Thesis for the Master's degree in Molecular Biosciences
Main field of study in Molecular Biology

60 study points

Construction of a vector system for functional verification of microRNA binding sites.

Erle Solheim

Department of Molecular Biosciences
Faculty of mathematics and natural sciences

UNIVERSITY OF OSLO

June 2009

Acknowledgements

The work presented in this thesis was carried out at the institute of Medical Genetics, Oslo University Hospital, Ullevål, from June 2008 to June 2009.

I would like to thank:

- My supervisor, Associate professor Eirik Frengen, for his mentoring, inspiration and help during the practical work and in the process of writing this thesis.
- My internal supervisor, Professor Tom Kristensen, for valuable feedback during the final steps of the writing process.
- Mona Mari Lindeberg for teaching and guiding me through my first months at the lab, and the rest of the Frengen group for taking an interest in my work and helping and guiding me whenever I needed it. Also, thank you for providing an including and inspiring work environment.
- Espen Enerly at the department of Genetics at the Radium Hospital, for the collaboration.
- Hans Christian Dalsbotten Aass at the core facility for flow cytometry at Ullevål, for analysing my samples and learning me all there is to know about flow cytometry. I must also thank Kirsti Solberg Landsverk at the core facility for flow cytometry at the Radium Hospital, for valuable feedback during the gating process.
- Hans Christian Åsheim at the department of Medical Genetics at Ullevål, for taking an interest in my work and giving me feedback on my results in the lab.
- My parents, my sister and my grandmother for support. Kari Skovlund for providing a place for us to live during the last year of my education. My boyfriend Johnny Long Skovlund for love and support during my five years at the University.

Abstract

MicroRNAs (miRNAs) are a relatively newly discovered class of small endogenous non coding RNA molecules with important roles in cell function and cancer development. Mature miRNAs bind to complementary sequences of transcripts and either cleave or inhibit translation of the mRNA. Some studies have indicated that grouping of tumor samples based on miRNA expression profiling correlate better than mRNA expression with cell differentiation and development. Experiments with diagnostic profiling using miRNAs have lead to suggestions for the use of miRNAs in cancer therapy. The properties of the binding between the target mRNA sequence and the miRNA however, are complicated and diverse and available computer programs designed to predict miRNA targets uses different algorithms and often disagrees in their predictions. This emphasizes the need for experimental verification of putative binding sites, predicted by such algorithms. The aim of this thesis was to construct a vector system for functional miRNA target verification with *EGFP* as the reporter gene. Flow cytometry was chosen as the method for analysing the EGFP expression of transfected cells. The results verified the construction of five vectors: the constructed pmiRPG and four constructs with the 3'UTR sequence from *KRAS, WWOX, CXXC4* and *WDR79* inserted in a fusion transcript with the *EGFP* gene. The results presented may indicate that a promoter exchange from a phosphoglycerate (PGK) promoter to a stronger promoter could enhance the fluorescence intensity of transfected cells and give better recordings. The flow cytometry analysis in the current thesis may indicate that the miRNA hsa-let-7a regulates the *KRAS* gene and that hsa-miR-487 regulates the *WWOX* gene. However an optimalization of the miRNA target verification system should be performed before a firm conclusion can be made concerning this putative miRNA-based regulation.

Table of contents

Abrevations

AMO	Anti-miRNA oligo nucleotides
BSA	Bovine serum albumin
CIP	Calf intestinal alkaline phosphatase
CMV	Cytomegalovirus
DMSO	Dimethyl sulfoxide
EDTA	Ethylene diaminetetraacetic acid
EGFP	Enhanced Green Fluorescent Protein
FITC	Fluorescein isothiocynate
FSC	Forward scatter
HEK cell	Human epithelial kidney cell
LB broth	Luria Bertani broth
LD	Loading dye
MCS	Multiple cloning site
miRNA	MicroRNA
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PGK	Phosphoglycerate kinase
PI	Propidium iodide
RISC	RNA-induced silencing complex
RPM	Rotations per minute
SNP	Single nucleotide polymorphism
SSC	Side Scatter
TAE	Tris-acetate EDTA
TSG	Tumor suppressor gene
TU	Transcription unit
U	Unit
UTR	Untranslated region
UV	Ultra violet

1 Introduction

1.1 MicroRNA

MicroRNAs (miRNAs) are a class of small (~18-25 nucleotides) endogenous non coding RNA molecules that regulate a wide array of developmental and physiological processes (reviewed by Tong *et al*. (2008) and Mirnezami *et al* (2009)). One of the first miRNAs discovered, lin-4, initially termed an stRNA (short temporal RNA), was found in *Caenorhabditis elegans*, and was revealed to play an important role in the processes in the transition from larva to adult stage (Lee et al., 1993). A second miRNA, Let-7 first identified in *C.elegans* as a temporal switch between larval and adult fates (Reinhart et al., 2000), was subsequently identified in a large number of species, including vertebrates. Thousands of miRNAs have now been identified and 851 human miRNAs have been registered in the online database, miRBase (Welcome Trust Sanger Institute: http://microrna.sanger.ac.uk/, Release 13.0, March 2009) that gathers miRNA sequences (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008). These miRNAs are predicted to regulate the expression of over one third of all human genes (Lewis et al., 2005) and have been shown to be involved in a variety of biological processes such as developmental timing, cell death, cell proliferation and patterning of the nervous system (Bartel, 2004).

miRNA genes are scattered across all chromosomes in humans except for the Y chromosome and approximately half of all known miRNAs is found in polycistronic clusters of often related, but sometimes unrelated miRNAs (Kim and Nam, 2006). Analyses of miRNA gene locations have showed that the majority of mammalian miRNA genes are located in defined transcription units (TUs) and can be categorized based on their genomic locations in introns of protein coding TUs, introns of noncoding TUs or exons of noncoding RNAs. An additional group are ‘mixed’ miRNA genes present in either exons or introns depending on the alternative mRNA splicing pattern (Rodriguez et al., 2004).

miRNAs are transcribed by the RNA polymerase II enzyme as long primary transcripts (pri-miRNAs) (Tong and Nemunaitis, 2008). The transcript is capped and poly adenylated in the

5’ and 3’ end, respectively, and folded into one or more hairpin structures. The hairpin is processed (Figure 1-1) by a complex composed of the nuclear RNase III enzyme Drosha and the RNA binding cofactor Pasha/DGCR8 (Tang, 2005). The structure called precursor miRNA (pre-miR) is transported out of the nucleus and into the cytosol by exportin 5 and its cofactor Ran-GTP, and further processed by a second RNase enzyme, Dicer, which crop the hairpin into a short miRNA duplex (Tang, 2005). A helicase will unwind this duplex (Salzman et al., 2007) before it is cleaved into a mature miRNA and incorporated into an RNA-induced silencing complex (RISC) (Meister and Tuschl, 2004). The miRISC effector complex, containing an Argonaut protein as the catalytic component (Tang, 2005), is guided to complementary sequences in mRNA transcripts and regulate post transcriptional gene expression by translational repression of mRNAs that bind to miRNA with imperfect complementarity, or by cleavage of mRNA targets that bind to miRNA with perfect sequence complementarity (Tong and Nemunaitis, 2008). Translational repression trough sites in 3’UTRs (untranslated regions) is considered to be the predominant mechanism in mammalian cells (Brennecke et al., 2005; Sethupathy et al., 2006; Tang, 2005).

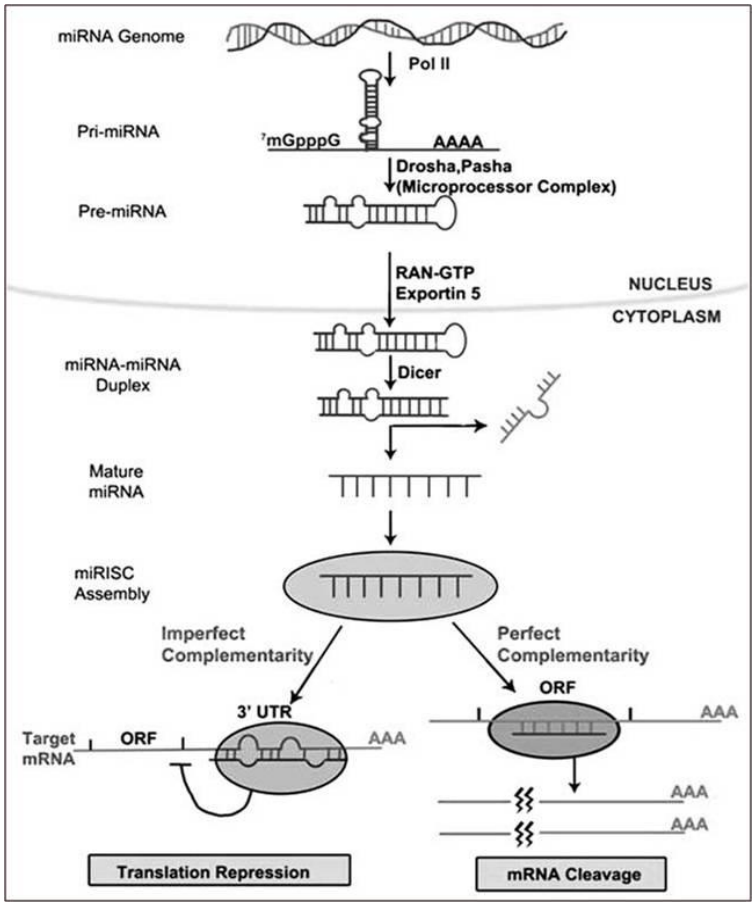


Figure 1-1: miRNA processing. **Nucleus**: A primary miRNA (pri-miRNA) is transcribed from DNA, processed and folded into the pri-miRNA hairpin structure which is cleaved by a microprocessor complex. The shorter stem loop, pre-miRNA, is transported across the nuclear membrane. **Cytoplasm**: The pre-miRNA is further processed by Dicer, cropped into a short miRNA duplex and unwound. After cleavage, the mature miRNA are incorporated into an RNA-induced silencing complex (RISC). The miRNA-RISC complex negatively regulates post transcriptional gene expression by translational repression or cleavage of the target mRNA dependent on the degree of complementarity. Figure from Tong *et al.* (2008).

1.2 miRNA in cancer

1.2.1 miRNA as oncogenes or tumor suppressors

Cancer arises from a stepwise accumulation of genetic alterations that drive the progressive transformation of normal human cells into highly malignant derivatives (Figure 1-2). The process has several steps that result in those alterations in cell physiology that lead to malignant cell growth, such as self sufficiency in growth signals, insensitivity to anti-growth signals, evasion from apoptosis, unlimited replicative potential, sustained angiogenesis, and tissue invasion and metastasis (Hanahan and Weinberg, 2000).

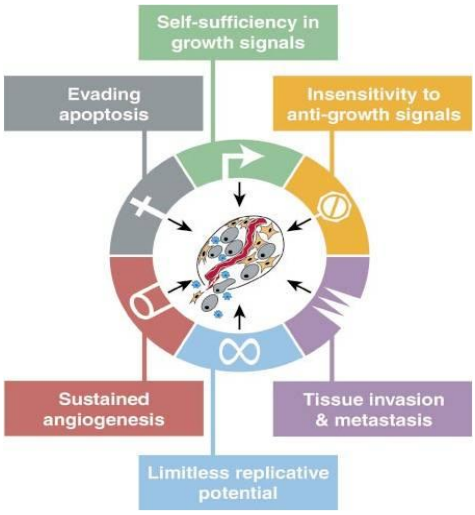


Figure 1-2: The six abilities indicated are suggested to be essential in the transformation from normal to malignant cells. The abilities may be acquired in various chronological orders in order to become cancerous. Figure from Hanahan *et al.* (2000).

miRNA expression correlates with various cancers, and some miRNAs are thought to function as tumor suppressors or oncogenes (reviewed by Esquela-Kerscher (2006)). One of the first examples of miRNA involvement in human cancer came from molecular studies characterizing the 13q14 deletion in human chronic lymphocytic leukaemia (CLL), which revealed that two miRNAs, miR-15a and miR-16-1, were the only genes within the smallest common region of deletion. The same miRNAs were found down regulated in over half of the human CLL-patients in one study (Calin et al., 2002).

Studies show that neoplasies have distinct miRNA signatures that are tissue specific and differ from those of the normal tissue counterpart (Bottoni et al., 2007; Iorio et al., 2005; Liu et al., 2004; Lu et al., 2005; Volinia et al., 2006; Yanaihara et al., 2006). In most cases, deregulation consistently acts in one direction, either up regulating or down regulating the

miRNAs, which may suggest that these miRNAs play an important role in tumorigenesis (Negrini et al., 2007).

miRNAs functioning as oncogenes are expected to have increased expression in tumors. These oncogene miRNAs usually promote tumor development by inhibiting tumor suppressor genes and genes that control cell differentiation and apoptosis. Many miRNA genes have been found that are significantly over-expressed in different cancers, and a large fraction of them appear to function as oncogenes (Zhang et al., 2007). The seven miRNAs on the polycistronic miR-17-92 cluster (miR-17-5p, miR-17-3p, miR-18a, miR-19a, miR-19b-1 miR-20a and miR-92-1) are examples of oncogenic miRNAs, and are located on chromosome 13q31, a locus that is amplified in several kinds of lymphomas and solid tumors (Hayashita et al., 2005; Ota et al., 2004). Compared to normal tissue, the expression of the miR-17-92 cluster is significantly increased in aggressive forms of lymphomas, small-cell lung cancer and human B-cell lymphoma, and it also appears to enhance lung cancer growth (Hayashita et al., 2005). Bioinformatic studies have indicated that numerous genes are targets of the miRNAs in the mir-17-19 cluster (Krek et al., 2005; Lewis et al., 2005). Examples are the tumor suppressor genes *PTEN* and *RB2* (Lewis et al., 2003) and the cell cycle transcription factor E2F1(O'Donnell et al., 2005).

miRNAs that are considered as tumor suppressor genes (TSG) often show decreased expression in some cancerous cells. Tumor suppressor miRNAs usually prevent tumor development by inhibition of oncogenes and genes that control cell differentiation or apoptosis (Zhang et al., 2007). Several miRNAs are considered as TSG, for example miRNA let-7, one of the first miRNAs to be discovered (Bartel, 2004). Inappropriate expression of let-7 results in oncogenic loss off differentiation, and in humans, let-7 is located on a fragile chromosome region that usually are deleted in human cancers (Calin et al., 2004). Studies have shown that the RAS family of proteins which acts as regulators of cellular growth and diffrentiation, are direct targets of the let-7 family. Let-7 binds to the 3'UTR of RAS and negatively regulates the gene expression by translational repression (Johnson et al., 2005).

1.2.2 miRNA in diagnostic profiling

Grouping of tumor samples based on miRNA profiling has been shown give a better correlation than mRNA expression profiling, with cell differentiation and development (Iorio et al., 2008) and miRNA microarray analysis and northern blotting indicates that miRNA expression profiles are good indicators for distinguishing even poorly differentiated tumour tissues apart from normal tissues (Lu et al., 2005). These findings suggest that miRNAs can be used as biomarkers for malignancy in humans and could be powerful diagnostic tools to classify tumors and to predict outcome and prognosis.

Various methods have been developed to classify tumors with regard to their miRNA expression, and to investigate the link between miRNA expression and cancer development and prognosis in humans. One of the first breast cancer specific miRNA signatures was obtained after a genome wide miRNA expression analysis on a large set of normal and tumor breast tissues (Iorio et al., 2005). In the same study, there were also identified miRNAs whose expression correlated with specific cancer biopathologic features such as expression of estrogen and progesterone receptor, tumor stage, vascular invasion or proliferation index (Iorio et al., 2005).

Microarray assays using the entire available miRNAome have been developed and are widely used for miRNA expression analysis (Liu et al., 2004; Volinia et al., 2006). A bead based flow cytometry technique has also been used to evaluate the miRNA profiling in different tumor subtypes (Lu et al., 2005). Metastatic cancer of unknown primary origin is often a very aggressive disease with poor prognosis (Pimiento et al., 2007) and Rosenfeld et al.(2008) developed a miRNA based tissue classifier system as a tool to classify the origin of tumor samples. The system was based on 48 distinctly expressed miRNAs and the classifier algorithm was constructed as a branched binary tree that was able to distinguish between 25 different types of cancers (Rosenfeld et al., 2008).

miRNAs can predict the clinical behaviour of certain cancers. Reduced expression of let-7 family members has, as an example been shown to correlate with reduced survival in lung cancer, and also shorter overall survival after curative surgery, independent of the disease stage (Mirnezami et al., 2009; Takamizawa et al., 2004). Other studies have found a

relationship between overexpression of miR-155 and poor post operative survival in lung cancer and B-cell lymphomas, suggesting that miR-155 may act as a predictor for aggressive tumor phenotypes (Calin et al., 2005; Yanaihara et al., 2006).

1.2.3 miRNA in future cancer therapy

Synthetic anti-sense oligonucleotides that encode sequences that are complementary to pri-, pre- or mature- miRNAs are called AMOs (anti-miRNA oligonucleotides). AMOs might be used to inactivate oncogenic miRNAs in tumors and slow down tumor growth (Weiler et al., 2006). AMOs that are conjugated with cholesterol, called "antagomirs", have been shown to effectively inhibit endogenous miRNA activity in various organs when injected into mice, and might be promising therapeutic agents (Krutzfeldt et al., 2005).

miRNA inactivation could alternatively be accomplished through frequent or continuous delivery of antisense oligonucleotides such as 2'-O-methyl- or locked nucleic acid (LNA modified nucleotides), which are specially designed to be more stable and less toxic than other cancer treatments in order to target transforming miRNAs such as miR-155 (Meister et al., 2004).

Transient expression systems that use viral or liposomal delivery might be useful for *in vivo* administering of large quantities of miRNAs such as the tumor suppressor family let-7. These techniques would involve expression of the pre-miRNA hairpin and flanking sequences, potentially under the control of tissue specific promoters (Esquela-Kerscher and Slack, 2006).

miRNAs could likely influence the response to chemotherapy or targeted therapies such as Trastuzumab, a monoclonal antibody against HER2, or anti-estrogens such as Tamoxifen by making the cells more vulnerable. miR-21, which directly targets the oncosuppressor PTEN, has been demonstrated to influence the response to chemotherapy, in particular gemcitabine (Meng et al., 2006). The let-7 family can suppress the resistance to anticancer radiation therapy and thereby alter the sensitivity to radiation (Weidhaas et al., 2007).

1.3 miRNA-mRNA interaction

1.3.1 miRNA target binding

miRNA target sites in the 3'UTR sequences of mammalian mRNAs can be divided into three categories with different binding properties (Figure 1-3). Targets with "5' dominant" sites base pair well to the 5'end of the miRNA binding site. Subtypes of this group are the "canonical" sites which pair well at both 3' and 5' ends of miRNAs, and "seed" sites which require little or no 3' end pairing support. Target sequences in the second group, "3' compensatory" sites have weak 5' base pairing and depend on strong compensatory binding in the 3' end of the miRNA (Brennecke et al., 2005).

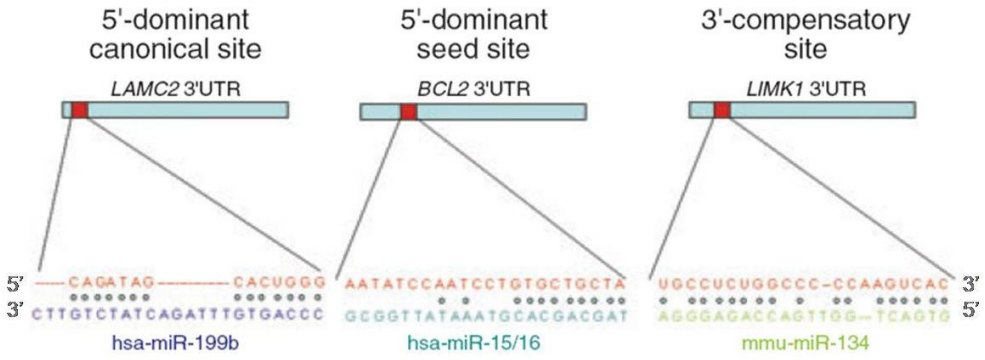


Figure 1-3: Three categories of miRNA target sites. Experimentally supported examples of mammalian miRNA target sites; canonical (left), seed (middle) and 3' compensatory (right). Figure from Sethupathy *et al*. (2006).

The most important element in miRNA binding is the "seed" sequence in the 5' region of the miRNA, centred on nucleotides 2-7, that bind to the mRNA sequence with Watson-Crick pairing. Use of seed match as a criterion for miRNA target site search, markedly reduces the occurrence of false positive predictions (Brennecke et al., 2005; Krek et al., 2005; Lewis et al., 2003) and improves prediction reliability (Lewis et al., 2003). Experiments have also showed that miRNA-like regulation was most sensitive to nucleotide substitutions that disrupted seed pairing (Brennecke et al., 2005).

In addition to the seed binding, there are numerous potential pairing possibilities involving the 3'end of the miRNA and the UTR. The "3'supplementary" sequences optimally centres on miRNA nucleotides 13-16 and bind to the UTR region directly opposite of this segment.

The association is sensitive to pairing geometry, and prefers at least 3-4 contiguous Watson-Crick pairs uninterrupted by bulges, mismatches or wobbles. This type of binding however, is atypical and plays only a modest role in target recognition (Grimson et al., 2007).

"3'compensatory" sequences of miRNAs can compensate for a single nucleotide bulge or mismatch in the seed region. The pairing of the 3' compensatory sequences centred on miRNA nucleotides 13-17 extends to at least nine contiguous Watson-Crick pairs. Extensive pairing to the 3'end of the miRNA however, is not sufficient to confer regulation on its own without a minimal element of complementarity in the 5'end (Brennecke et al., 2005).

The efficiency in which the miRNAs bind to the target mRNAs ranges from efficient binding with an 8mer seed sequence to marginal binding with a 6mers seed. Dual binding of the seed sequence in combination with compensatory or supplementary binding such as a 7mer-8mer or two 7mers are also considered efficient (Bartel, 2009).

A "seed nucleation" model (Figure 1-4) proposes how binding of the miRNA to the mRNA 3'UTR is nucleated at the seed match, spreads to the central and 3' regions of the miRNA and leads to repression of translation ,or less common cleavage of miRNA (Bartel, 2004, 2009). miRNAs bound to Argonaut proteins are organized to favour recognition and pairing to 8mer sites in the mRNA, by a binding pocket. An A at position 1 of the site is recognized directly by proteins of the silencing complex. Massive conformational accommodation (Figure 1-4c) of extensive paired sites leads to pairing, anchoring and nucleation at the seed. The seed nucleation extends to the central region of the miRNA and causes the protein to loosen its grip on the 3'region of the miRNA, and thereby allows the miRNA and mRNA to wrap around each other. This pairing can lead to mRNA cleavage when the Argonaut protein locks down on the paired duplex. 3'supplementary pairing can lead to translational repression (Figure 1-4e) when the mRNA pairs to nucleotides 13-16 and incorporates them into a short helical segment. In this mode of recognition however, the miRNA and mRNA are not wrapped around each other (Bartel, 2009).

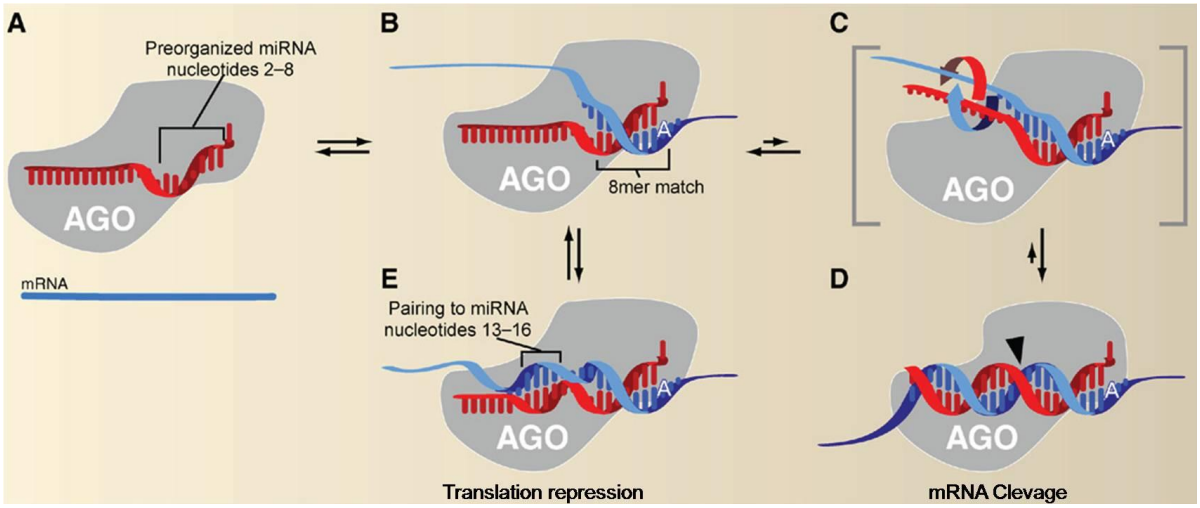


Figure 1-4: **Speculative "Seed nucleation" model**. **a)** miRNA (red) bound to Argonaut (AGO) protein (grey) favour pairing to an incoming mRNA strand (blue). **b)** Recognition of 8mer site by the binding pocket **c)** Conformational changes cause the protein to loosen its grip on the 3'region of the miRNA, allowing the miRNA an mRNA to wrap around each other. **d)** The Argonaut protein locks down on the paired duplex, which places the active site in position (arrow head) to cleave the mRNA. **e)** The mRNA can pair to nucleotides 13-16, incorporating them into a short helical segment in a 3'supplementary pairing which leads to translation repression. Figure from Bartel *et al*. (2009).

The location of sites within the 3'UTR sequence can have a positive influence and enhance the effect of miRNA binding. The position of the miRNA binding site in the 3'UTR is reported to be more conserved and more effective when located 15 nucleotides or more after the stop codon (Grimson et al., 2007) and a genome wide analysis of site conservation, site efficiency and site depletion indicated that sites tend to be most effective if they do not fall in the middle of a long UTR sequence (Grimson et al., 2007). The same analysis also showed that the nucleotides in the immediate vicinity of the site are important, and sites within a region of high local AU content were performing best. The authors discovered that sites close together (within 40nt and minimum 8nt apart) tend to act cooperatively, leading to enhanced repression over that expected from two separate sites (Grimson et al., 2007).

miRNA activity can be classified according to different effects on the protein level, phenotype or number of target sites, and the regulatory effects on the target transcript are diverse and complex (reviewed in Flynt (2008)). A summary is presented in Figure 1-5. When assaying at the protein level, the "switch targets" are turned off by miRNAs. "Tuning targets" in contrast remain functional but reduced in quality, and neutral targets have no consequence on the protein level. The intrinsic detectable miRNA contribution to morphology, physiology or behaviour is called genetic switches and is detected as a change

in phenotype. miRNA regulation can also be classified according to whether a major effect is mediated through one or a few target sites or many targets (target battery). All known switches concern cases of one or a few important miRNA target sites (Flynt and Lai, 2008).

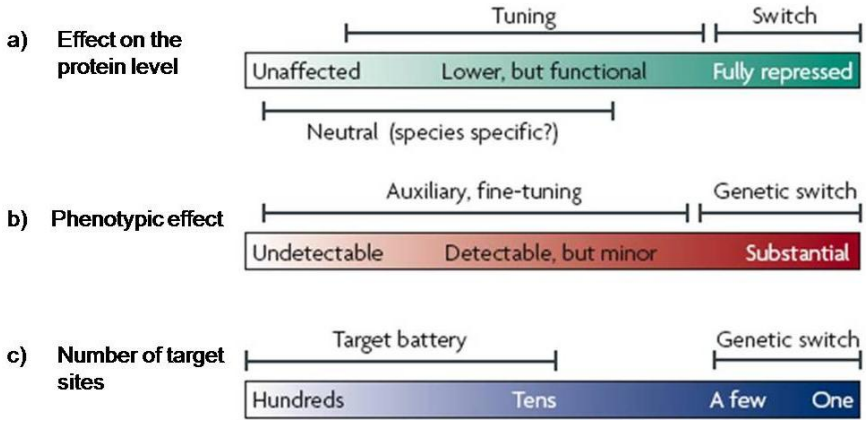


Figure 1-5: Classification of miRNA activity and function. **a)** Targets that become inactive following miRNA mediated repression are called "Switches", whereas "tuning" targets continue to produce functional proteins in reduced amounts. Neutral targets have no consequence for the cell, and are possibly species specific. **b)** Targets with genetically defined and phenotypic consequences are called "miRNA genetic switches". Other targets only have fine tuning or auxiliary effects **c)** The total number of targets who are repressed by miRNA can be detected quantitatively for a given effect. Genetic switches are regulated through only one or a few targets. Detectable effects can also originate from the combination of hundreds of regulated targets. Figure from Flynt *et al*. (2008).

1.3.2 Identification of miRNA targets

The parameters that are important for efficient miRNA target recognition and specificity are still not fully understood, and the available tools for miRNA target predictions do not always agree on which miRNAs that bind to a given target sequence.

Watanabe *et al*. (2007) presented a schematic summery in the form of a pipeline of common algorithms used in the prediction of miRNA targets (Figure 1-6). The first step is the extraction of orthologous miRNA and mRNA 3'UTR sequences. The base pairing between miRNAs and their targets can be analyzed next, to check for features such as stable binding at the 5' end of the miRNA. Thermodynamic analysis of miRNA-mRNA duplexes is performed calculating the free energy of the duplex formation and binding. Cross species sequence comparison is used to investigate the evolutionary conservation between related species. The number of target site for the miRNA can be counted as a final step, since mRNAs are likely to be regulated by miRNA binding to more than one target sites

(Rajewsky, 2006; Watanabe et al., 2007). The programs used for miRNA target predictions in this thesis used variations of this schematic pipeline, with different main focuses as discussed below.

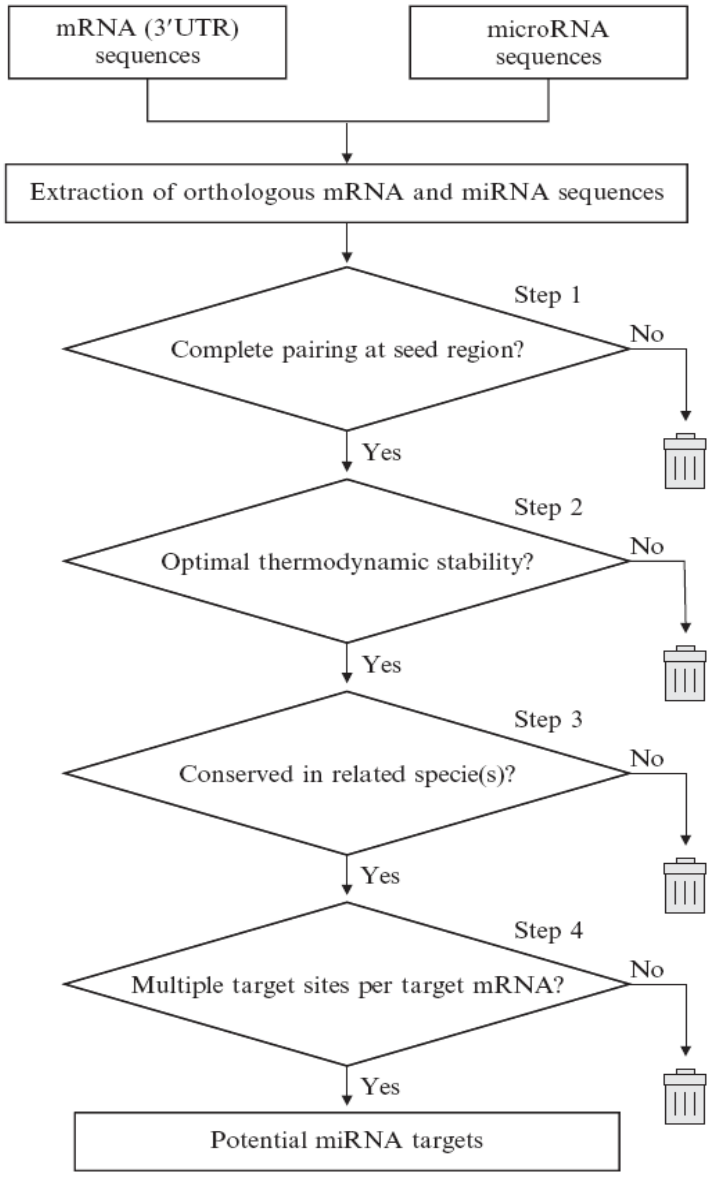


Figure 1-6: A schematic pipeline for the main steps in miRNA target predictions: Datasets from miRNAs and mRNAs (3'UTR) are provided for input and similar datasets from related species are constructed using data on putative orthologs. Preparations of the datasets leads to the identification of miRNA binding sites through four steps: **1)** Determination of base pair patterns according to the complementarity within specific regions. **2)** Determination of the strength of the miRNA-mRNA duplex by calculation of the free energy. **3)** Comparative analysis. **4)** Check for the presence of multiple target sites per mRNA transcript. Figure from *Watanabe et al*. (2007)

The miRNA registry, miRBase (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008) includes information about the potential genomic targets of miRNAs in addition to the miRNA sequence data. The miRBase Target database thereby provides predicted miRNA target transcripts for various species with the use of the database and the miRanda algorithm. The miRanda software was initially designed to predict miRNA targets

in *Drosophila melanogaster* but has also been applied to predict human miRNA targets (Enright et al., 2003; John et al., 2004). The algorithm uses three steps starting with the identification of a potential target sequence. The free energy is calculated for the miRNA-mRNA duplex and in the final step, the evolutionary conservation are established (Enright et al., 2003; John et al., 2004).

TargetScan is an algorithm developed for prediction of miRNA targets in vertebrates (Lewis et al., 2003) and the TargetScanS is an improved algorithm which requires a shorter seed match (six nucleotides) and is independent of thermodynamic stability or multiple target sites. The algorithm requires the presence of conserved adenosines around the seed sequence (Lewis et al., 2005).

The DIANA-microT computational program uses the necessity of a central bulge and strong binding at the 3' end of the miRNA, when the 5' seed pairing is weak. This method uncovers predominant miRNA targets that contain only single target sites (Kiriakidou et al., 2004). The PicTar software fully relies on comparative data from several species to identify common targets for miRNAs (Krek et al., 2005). The likelihood that a given sequence is bound by one or more miRNAs is estimated and the target genes are predicted using a common criteria, such as optimal binding free energy. The site is then tested statistically using genome wide alignment of eight vertebrate genomes to filter false positives (Krek et al., 2005).

In this study the programs showed a wide variation of miRNA target predictions. The few times they agreed, there were seldom more than two programs that predicted the same miRNA for a specific gene. This emphasizes the need for experimental verification of putative miRNA binding sites. The approach chosen in the present study was to select miRNAs that were predicted by two or more programs to target a chosen transcript, and use them in cotransfections with a reporter vector. The miRNAs selected for each of the genes are listed in Table 3-3.

1.4 Project background

1.4.1 Choice of method

There are several vectors and assays available for miRNA target analysis, and the common feature of these systems is the use of a gene encoding a luciferase protein that produce light through a chemical reaction (Marques and Esteves da Silva, 2009). The light produced and recorded directly from the whole cell population is used as a measure of the level of gene expression. The values represent the average intensity of all cells.

In this thesis the idea was to use enhanced green fluorescent protein (EGFP) a mutated (Heim et al., 1995; Patterson et al., 1997) version of GFP from the jellyfish *Aqueorea Victoria* (Prasher et al., 1992) as the reporter gene in the expression vector. EGFP emits green light (fluorescence) when exposed to ultra violet light, and can be analysed by flow cytometry. The advantage using this approach is that it is possible to discriminate between transfected and nontransfected cells, and to record and compare fluorescence intensity from single cells in the transfected population of cells.

To our current knowledge, there are no vectors available for miRNA target analysis that use *EGFP* as the reporter gene, and consequently, an expression vector for this thesis using EGFP had to be constructed.

1.4.2 Genes for cloning into the reporter vector

1.4.2.1KRAS

KRAS (GenBank: M54968) is one of the founding members of the *RAS* (rat sarcoma) superfamily of small GTPases (reviewed in Quinlan *et al*. (2009)). This family of proteins contains master regulators of a wide range of cellular processes including growth, proliferation, survival, apoptosis, senescence, cytoskeletal dynamics, vesicular trafficking, differentiation and gene expression (Giehl, 2005; Malumbres and Barbacid, 2003; Rajalingam et al., 2007).

The *RAS* genes are frequently mutated in human cancer and it is estimated that approximately 15-20% of all human tumors have activating mutations in one of the *RAS* oncogenes or components of the Ras signalling pathway. Several approaches are made to develop tumor therapies effectively targeting RAS and RAS effector pathways (Adjei, 2001; Downward, 2003).

The family of *RAS* genes has been shown to be regulated by miRNAs and the let-7 miRNA has a documented strong effect on the regulation of the *KRAS* gene. Let-7 binds to the 3'UTR sequence of *KRAS* mRNA and effectively reduces the expression of KRAS (Johnson et al., 2005). The *KRAS* 3'UTR sequence was included in this thesis primarily as a positive control that would be valuable for the verification of the functionality of the miRNA binding site reporter system.

1.4.2.2 CXXC4

CXXC finger 4 (*CXXC4*, GenBank: NM_025212) encodes the protein called Idax (Inhibition of the Dvl and Axin complex) which acts as a negative regulator of the Wnt signalling pathway by interacting with Dishevelled (Dvl)(Hino et al., 2001).

Kojima et al. (2009)showed that CXXC4 mRNA levels in tumor samples were significantly lower in patients with metastases compared to those without, and in addition, patients with poor cause-specific survival outcome showed decreased CXXC4 expression. Decreased CXXC4 expression has been linked to aggressive renal cell carcinoma (RCC) (Kojima et al., 2009). Knockdown experiments have revealed that the lowered CXXC4 expression leads to activation of the Wnt/ β-catenin signalling and thereby promotes cell proliferation and survival in RCC cells. The 3'UTR sequence from this gene was chosen for study in this thesis because miRNAs may be a factor involved in the downregulation of this gene.

1.4.2.3 WDR79

WD repeat domain 79 (WDR79, GenBank: NM_018081), also called WRAP53 (WD repeat containing, antisense to TP53) codes for a protein that belongs to the large WD-repeat protein family with diverse functions (Smith et al., 1999) and has recently been associated with Cajal body-specific RNPs (scaRNPs) and with telomerase RNA (Tycowski et al., 2009)

WDR79 is transcribed in an antisense orientation relative to the tumor suppressor gene TP53, thus juxtaposing proximal promoter elements between these two genes. The WDR79 transcript could include regulatory elements for the transcription of TP53 mRNA (Laitinen et al., 2004)

Garcia-Closas et al. (2007) revealed a linkage disequilibrium (LD) between TP53 and its flanking gene WDR79, and their data suggested that common variants in regulatory regions of TP53 and WDR79, could be related to increased risk for development of ER negative breast cancer (Garcia-Closas et al., 2007). This gene is interesting due to its relation to the widely known tumor suppressor gene TP53, and was therefore chosen for studies of potential miRNA binding to the 3'UTR sequence.

1.4.2.4 WWOX

The WW domain-containing oxidoreductase gene (*WWOX*, GenBank: AF211943) contains nine exons, and encodes an mRNA that is 2.2kb long. The gene is located on chromosome 16q 23.3-24.1 in a region recognized as the fragile site, FRA16D (Bednarek et al., 2000).

WWOX might play a central role in different signal transduction pathways and a number of cellular processes including transcriptional repression, growth and apoptosis (Aqeilan and Croce, 2007). Studies have shown that *WWOX* expression is up regulated in endocrine organs such as testis, ovary and breast, indicating the importance of *WWOX* in these tissues (Bednarek et al., 2000).

Studies demonstrate that *WWOX* expression is lost or reduced in a variety of human malignancies making *WWOX* the fragile gene target of genomic alterations at 16q23 (Ludes-Meyers et al., 2003). The frequent deletions of *WWOX* in multiple tumors suggest that it may act as a tumor suppressor gene involved in a variety of cancers like breast, ovarian, prostate, esophageal, lung, gastric and hepatic (Hezova et al., 2007). Over-expression studies indicates that the gene might be capable of suppressing tumor growth, and ectopic expression of *WWOX* has been shown to strongly inhibit anchorage independent growth of several breast cancer cell lines (Bednarek et al., 2000).

In a mutation screen of the coding region of the *WWOX* gene in a breast cancer cell panel, no evidence of mutation within the nine exons of the *WWOX* gene was detected (Bednarek et al., 2000; Driouch et al., 2002).Thus the mechanisms responsible for the downregulation of *WWOX* in cancers are unclear. In this thesis we aim at testing if candidate miRNAs regulate WWOX expression through binding to the *WWOX* 3'UTR.

1.5 Aims of the thesis

In the present study, the aim was to construct a reporter vector system for functional verification of miRNA binding sites. The vector includes *EGFP* as the reporter gene. The following aims were pursued in this thesis.

1. Construction of the vector.
2. Cloning of the 3'UTR sequences of the selected genes into the vector.
3. Use of online miRNA target prediction programs to predict miRNAs for the selected genes.
4. Use of flow cytometry to analyse changes in the EGFP expression levels when a human cell line is cotransfected with the vectors and miRNAs.

2 Material and methods

This chapter presents materials and methods used in the thesis. Solutions were made according to Sambrook *et al*.(2001), if nothing else is specified.

2.1 Microbiological techniques

2.1.1 Medium and Agar preparations

LB medium and LB agar were made according to standard protocols (Sambrook and Russel, 2001), and autoclaved immediately after preparation to achieve sterilization and to dissolve the solid reagents. The media and agar were stored at 4°C and preheated to 37°C before use with *E.coli* bacteria.

Selective LB media and LB agar were made to select for bacteria transformed with plasmids containing a gene for antibiotic resistance.

LB agar plates were made using a microwave oven to melt the agar. The agar was cooled down to approximately 50°C. When acquired for selection, ampicillin (Sigma-Aldrich, St. Louis, MO, USA) was added to a final concentration of 100µg/ml and 15ml melted LB agar was transferred to each Petri dish and cooled down to solidify. The LB agar plates were stored at 4°C.

2.1.2 Growth and storage

The *E.coli* strain DH5α (Invitrogen, Carlsbad, CA, USA) was used for cloning experiments and for general selection and amplification of plasmid DNA. Transformed bacteria were streaked onto selective LB-agar plates and incubated overnight at 37°C to obtain single colonies. Colonies were inoculated in separate selective LB-medium and incubated at 37°C overnight with agitation of 225 rpm in a Minitron Incubator Shaker (Infors AG, Bottmingen,

Switzerland). For prolonged storage, glycerol stocks were made from overnight bacterial cultures mixed with glycerol to a final concentration of 15% and stored at -80°C.

2.1.3 Preparation of competent *E. coli*

Competent bacteria were made according to the protocol from Inoue *et al*. (1990). A frozen stock of E. coli DH5α™ (Invitrogen, Carlsbad, CA, USA) was thawed and streaked out on non-selective LB-Agar plates and incubated over night at 37°C. The next day, a few colonies were inoculated in 250 ml SOB-medium in a sterile Erlenmeyer flask. The bacteria were incubated at 30°C with vigorous shaking. A spectrophotometer was used to register cell density. Incubation was stopped when the cell density reached OD_{600} = 0.6. The flask was chilled on ice for 10min and the cells were harvested by centrifugation at 6000 x g for 10min at 4°C. The cells were resuspended in 80ml ice cold TRB (10 mM Pipes, 15 mM $CaCl_2$, 250 nM KCl, 55 nM $MgCl_2$), placed on ice for 10 min and centrifugated again under the same conditions. Next, the cells were resuspended in 20 ml TRB, and gently while stirring, Dimethyl sulfocside (DMSO) (Sigma-Aldrich, St. Louis, MO, USA) was added to a concentration of 7%. The cells were left on ice for 10 min. 500µl cell suspension was aliquoted in eppendorf tubes, and snap frozen in dry ice and ethanol. The competent cells were stored at -80°C.

2.1.4 Heat shock transformation of *E. coli*

Competent DH5α *E. coli* was transformed with plasmids containing genes for antibiotic resistance (Cohen et al., 1972) according to the protocol of Sambrook *et al*. (2001). A tube of 500µl competent cells was thawed on ice, and aliquots of 50 µl were transferred to eppendorf tubes on ice. 1-5µl DNA solution was added and mixed gently into the cells with a pipette tip. As a control, 250 pg of pUC19 (Invitrogen, Carlsbad, CA, USA) was added in a separate tube. The transformation mixtures were incubated on ice for 30 min, heat shocked at 42°C in 30 seconds, and placed back on ice for 2 min. The cells were transferred to 5ml tubes and added 950µl pre-heated (37°C) LB-media and incubated for one hour at 37°C, with agitation at 225 rpm in a Minitron Incubator Shaker (Infors AG, Bottmingen, Switzerland). The incubation allows the bacteria to recover and express the antibiotic marker encoded by the

plasmid. 100-500 µl of each transformation reaction was streaked on to pre-heated LB-Agar plates with 100µg/ml ampicillin (Sigma-Aldrich, St. Louis, MO, USA). The plates were incubated overnight at 37°C.

The transformation frequency was calculated from the number of transformants on the pUC19 control-plate, divided by the amount of DNA added, in µg.

2.2 DNA techniques

2.2.1 Isolation of plasmid DNA

Plasmids were isolated from *E. coli* using the QIAprep® Spin Miniprep Kit (Qiagen®, Hilden, Germany) according to the manufacturer's protocol. In short, bacterial pellets from 5ml overnight cultures were resuspended and lysed under alkaline conditions. Sodium dodecyl sulphate (SDS) causes precipitation of cell debris which traps the genomic DNA. Plasmids however remain suspended. The lysis reactions were neutralized and the supernatants were transferred to a provided column, where the plasmid DNA was absorbed onto a silica membrane. After necessary washing steps to remove endonucleases and salts, the DNA was eluted in 30-50µl elution buffer (Qiagen®, Hilden, Germany) for a total yield of up to 20µg of DNA (high copy plasmids). For bigger experiments QIAGEN Plasmid Midiprep Kit (Qiagen®, Hilden, Germany) was used to purify up to 100µg of DNA (high copy plasmids) from a 25ml culture. The method was similar to the Miniprep except for the last steps. After elution of the DNA from the column, the DNA was precipitated with isopropanol in combination with centrifugation. The resulting pellet was washed in 70% ethanol, air dried and resolved in 50 µl nuclease free water or TE-Buffer (Qiagen®, Hilden, Germany). Purified plasmids in solution were stored at -20°C or at 4°C for shorter periods.

2.2.2 DNA precipitation

Precipitation was used to purify DNA resolved in liquid (Sambrook and Russel, 2001). The solution containing DNA was mixed with 10% 3M Sodium Acetate and 2.5 times the

volume of 100% ethanol. The mixture was centrifugated at 15000 x g at 4°C for 30 minutes. The supernatant was removed, and an ample amount of 75% ethanol was added to wash the DNA properly. The tube was centrifugated again for 10 minutes. The supernatant was removed and the pellet was left to air-dry in room temperature. The dried DNA was resuspended in nuclease free water or Elution Buffer (Qiagen®, Hilden, Germany). To separate plasmids from oligo nucleotides, 0.6 volumes of isopropanol were used instead of ethanol to enrich for vector fragments relative to oligo nucleotides.

2.2.3 Purification of DNA fragments from an Agarose Gel

DNA fragments separated on an agarose gel (section 2.2.7) were cut out and purified using QIAqick® Gel Extraction Kit (Qiagen®, Hilden, Germany). In short, the gel slice containing DNA fragments of 70bp-10kb was incubated at 50˚C with 3x volumes of a supplied buffer to melt and dissolve the gel. When the solution was homogenous (after approximately 10min), it was applied to a QIAqick column (supplied) and spun down to bind the DNA in the column. The DNA in the column was washed an additional time with the same buffer to remove all gel traces, and finally washed and eluted in the same way as with the QIAprep® Miniprep Kit protocol (see 2.2.1:Isolation of plasmid DNA).

2.2.4 Quantification of DNA

Quantification of DNA was done using NanoDropTM 1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) according to the manufacturer's user manual. In short, the buffer in which the DNA was resolved was set as blank, and one µl of undiluted DNA sample was applied to the instrument. The sample absorbance of UV-light with wavelength of 260nm was detected, and the concentration (ng/µl) was calculated according to Beer's Law. To additionally assess the purity of the DNA sample, a ratio of sample absorbance at 260nm and 280nm was recorded. A 260/280-ratio of 1.8-1.85 was accepted as pure DNA. Values below may indicate the presence of proteins and values above may indicate presence of phenol or RNA.

2.2.5 Restriction enzyme digestion

Restriction digestion of DNA, for analytical purposes, was prepared in a total reaction volume of 10 µl. 200-400ng of DNA was digested by 5-10 units of restriction enzyme. One unit (U) is defined as the amount of enzyme that digests 1µg DNA in one hour at optimal conditions. Appropriate buffer (10x concentration) and bovine serum albumin (BSA) (100x) provided by New England Biolabs (Ipswich, MA, USA) were diluted to a 1x concentration, and nuclease free water was used to attain the correct reaction volume. For a preparative restriction cutting, 2 µg of DNA and 10-20U of enzyme were used. The volume was adjusted to avoid more than 10% of enzyme volume in the total reaction volume. The enzymatic reaction was left on a heating block for one hour at the optimal temperature for the enzyme used.

2.2.6 Dephosphorylation of DNA

Dephosphorylation reactions were used to remove the 5'-phosphate from linear DNA fragments. This was accomplished with use of the dimeric glycoprotein calf intestinal alkaline phosphatase (CIP) (NEB, Ipswich, MA, USA). Dephosphorylation was done to decrease incidences of vector re-circularization, in experiments where a vector and an insert sequence were digested with the same enzyme and ligated together. This is an efficient way to diminish vector background in cloning experiments. The CIP-reactions were carried out as recommended by the provider (NEB, Ipswich, MA, USA). Generally, 0.5U CIP enzyme was used per µg of DNA.

2.2.7 Agarose gel electroforesis

DNA fragments of 37bp-10kb were analyzed on an agarose gel using 1 X TAE buffer with 1-2% SeaKem® LE agarose (Cambrex Bio Science, Rockland, ME, USA) added. Buffer and agarose powder were boiled to obtain a homogenous solution. After heating, the solution was cooled down to approximately 60°C and the SybrSafe® DNA gel stain (Invitrogen, Carlsbad, CA, USA) was added in an amount that diluted the 10 000 x solution to 1 x. Typically 5 µl Sybr Safe was used in 50 ml TAE. The solution with DNA fragments was added Gel

Loading Dye (NEB, Ipswich, MA, USA) (provided in a 6x concentration) to a final concentration of 1x. The solution contained 2,5% Ficoll 400 to increase the density of the sample, and trace amounts of SDS to remove enzymes remained bound to the DNA. The dye also contained EDTA that binds magnesium to stop enzymatic reactions. Bromophenol blue was the tracking dye and it migrate equivalent to a 300bp fragment on a 1% agarose gel. This made it possible to follow the migration through the gel, and thereby estimate when to stop the gel. The DNA fragments were separated according to their size and migrated towards the positive pole. A ladder sample with known fragment sizes was also loaded on to the gel for analytical purposes (NEB, Ipswich, MA, USA; Fermentas, Burlington, Ontario, Canada). UV-light with 254nm was used to detect the DNA fragments after ended electrophoresis. To visualize the smallest fragments 37bp-100bp, additional gel stain was added to the buffer a few minutes before the electrophoresis was ended.

2.2.8 Ligation of DNA fragments

Digested DNA fragments were ligated together with the use of bacteriophage T4 DNA ligase as recommended by the supplier (Roche™, Basel, Switzerland). The enzyme catalyzes the formation of phosphodiester bond between the 5'phosphate end and the 3'hydroxyl end in double stranded DNA. Generally, 100ng of the linear vector was used, and the molar ratio between the vector and the insertion fragments was set from 1:3 to 1:20 depending on the amount of DNA left after all preparative steps (digestion reactions, precipitation). The reaction volume was kept as small as possible (maximum 20µl), and 1U of enzyme was used per reaction. The reaction was incubated overnight at 16°C (sticky ends).

2.2.9 Primer design

Primers used in PCR or sequencing reactions were designed using the web based tool Primer3 (http://frodo.wi.mit.edu/). The DNA sequence of interest was specified and the program made suggestions for one or more primer pairs flanking the sequence. The primer estimations were made with fixed parameters including the GC content, primer length and melting temperature to ensure the quality of the primers. The oligo nucleotides were

purchased from MWG Biotech (Ebersberg, Germany). All primers used in this thesis are listed in the Appendix, Table 1.

2.2.10 PCR Reaction

Polymerase chain reactions (PCR) were used to amplify specific DNA fragments. A thermo stable polymerase was used for automated reaction cycles (Saiki et al., 1988). For this thesis $Vent_R^{®}$ DNA polymerase (NEB, Ipswich, MA, USA) was used. This polymerase has high fidelity caused by the 3'-5' proofreading exonuclease activity (Mattila et al., 1991). The optimal temperature was 72°C and the enzyme was stable at 95°C, the temperature used to separate double stranded DNA. The melting temperature was set 1°C below the estimated melting temperature of the primers, and the DNA polymerase extension time was set to 30 seconds per 500bp. The program that was adjusted for each reaction was as follow: 2 min at 95°C, followed by 30 cycles with 95°C denaturation for 30 sec, X°C annealing for 30 sec and 72°C extension for Y seconds. The reaction was prepared in PCR-tubes in 50µl total reaction volume containing 1U of the polymerase, 1x Vent buffer (NEB, Ipswich, MA, USA), 0.01mM of each primer and 10mM of dNTPs. 10ng of DNA template was added and the reaction was run in a GeneAmp® PCR system 9700 (Applied Biosystems Inc, Foster City, CA, USA).

2.2.11 Sequencing of DNA

Sequencing was done to verify new DNA sequences in expression vectors. The method was carried out similar to the PCR reaction (2.2.10) with some important differences. The reactions were set up with forward and reverse primer in separate tubes. The nucleotide mix contained the four dNTP's and in addition a portion of ddNTP's labelled with four different fluorescent dyes. The dye labelled ddNTP's terminates the elongation and when they were incorporated randomly, multiple molecules of varying length were generated.

In this thesis all sequencing experiments were done using the primers *pmiRPG-5850-Fwr* and *pmiRPG-6136-Rev* (Appendix, Table 1). BigDye® Terminator v 3.1 Cycle sequencing kit (Applied Biosystems Inc, Foster City, CA, USA) was used to prepare the sequencing

reaction according to the manufacturers instruction. Automated cycles were performed by the GeneAmp® PCR system 9700 (Applied Biosystems Inc, Foster City, CA, USA) and started with 1min of 96°C, followed by 25 cycles of 96°C denaturation for 10 sec, 50°C annealing for 5 sec and 60°C extension for 4min. The sequencing reaction was washed using Montage SEQ_{96} Cleanup Kit (Millipore, Billerica MA, USA) and analyzed by the ABI 3730 DNA analyzer (Applied Biosystems Inc, Foster City, CA, USA).

The final sequencing results were analyzed using the Contig Express program in Vector NTI Advance™ 10 (Invitrogen, Carlsbad, CA, USA) and the NCBI search program Nucleotide Blast 2 Sequences (http://blast.ncbi.nlm.nih.gov/bl2seq/wblast2.cgi) was used to compare the sequences to the expected GenBank DNA sequences.

2.3 Mammalian cell techniques

2.3.1 Cell line

The HEK293T cell line used in this thesis was provided by American Type Tissue Culture (ATCC, Manassas, VA, USA). This cell line is virus transformed, and has human epithelial kidney origin. HEK293T cells were chosen for their properties of being easy to work with and highly transfectable using conventional techniques. In this thesis, HEK293T cells were used for transient transfection with expression vectors and microRNA oligo nucleotides supplied by Ambion (Applied Biosystems Inc, Foster City, CA, USA).

2.3.2 Cell culturing and storage

HEK293T cells were cultured in RPMI 1640 media with 25mM Hepes and L-glutamine (BioWhittaker®, Lonza, Verviers, Belgium) supplied with 10% GIBCO™ fetal bovine serum (FBS) (Invitrogen, Carlsbad, CA, USA). The cells were kept in an incubator with 5% CO_2 and 37˚C. Cell work was carried out in a vertical laminar flow bench using sterile techniques.

The cells were grown in sterile flasks (Nunc™, Rochester, NY, USA) and diluted when reaching about 90% confluence. The media was removed and the cells were washed gently twice with Dulbecco's PBS (PAA Laboratories, Pasching, Australia). The cells were detached from the flask surface with GIBCO™ Trypsin (Invitrogen, Carlsbad, CA, USA) at 37˚C for two minutes. Trypsin was inactivated with RPMI 1640 medium and the cells were harvested at 1000 x g for 5 minutes. The cells were diluted 1:20 or according to the next scheduled experiment. Generally, medium was changed every third day.

Cells at a low passage number were stored in liquid nitrogen over longer time periods. Cultured cells were diluted to an approximate concentration of 1-2 x 10^6 cells per ml and supplied with 10% of DMSO. The cells were frozen in cryo-tubes, stored at -80˚C overnight and transferred to a nitrogen tank.

2.3.3 Cell quantification

NucleoCounter (ChemoMetec A/S, Allerød, Denmark) was used for cell counting according to the manufacturer's instructions. Cells were lysed with a reagent A-100 which disrupts the plasma membrane in mammalian cells. A second reagent, stabilizing reagent B was used next to allow propidium iodide to stain DNA in the exposed nuclei. Equal amounts of the cell sample, reagent A and B were used. 50µl of the sample mixture was loaded into a nucleocasette and inserted into the NucleoCounter, which calculated the cell number based on the number of stained nuclei detected.

2.3.4 Transient transfection of cell line HEK293T

Transfection is a method used to introduce exogenous DNA into human cell lines. The method used utilizes lipids that form complexes with DNA. The lipid complexes enable import of foreign DNA into the cell, over the hydrophobic lipid membrane. Volumes and concentrations of reagents were as in the Lipofectamine™ 2000 protocol for DNA transfection (Invitrogen, Carlsbad, CA, USA). Master solutions were made to ensure equal distribution of Lipofectamine 2000™ (Invitrogen, Carlsbad, CA, USA) and DNA.

At day one, $1.6x10^5$ cells were plated out in each well in a 12 well tray. The second day, the transfection mixture was made. Lipofectamine™ 2000 (Invitrogen, Carlsbad, CA, USA) and appropriate amounts of DNA (1.6µg of vector and 1.56-100pmol of miRNA and SiRNA) were separately diluted in serum-free BioWhittaker® -media (Lonza, Verviers, Belgium). After five minutes of incubation, the DNA/media mixture was transferred to the tube containing the lipofectamine/media mixture and incubated at room temperature for 20 min. During the incubation, complexes of Lipofectamine 2000 and DNA were formed. The transfection mixtures were carefully dripped over the cells. The cells were incubated overnight at 37˚C and with 5% CO_2. On the third day the cells were harvested for flow cytometry or visualized under UV-light.

2.3.5 Cell harvest

Cells were harvested prior to a flow experiment, 24 hours after transfection. The media was carefully removed and the cells were added appropriate amounts of PBS (PAA Laboratories, Pasching, Australia). After detachment of the cells with a pipette, the cells were collected by centrifugation at 6000 x g, 4°C for 5min, and resuspended in 350µl PBS. The cells were transferred to FALCON® 352063 tubes (BD Biosciences, Franklin Lakes, NJ USA) and kept on ice until analysis by Flow Cytometry.

2.4 Flow cytometry

Flow cytometry was used in this thesis to measure fluorescence intensity in HEK293T cells transfected with plasmids containing a gene encoding enhanced green fluorescent protein (EGFP). In short, the harvested (2.3.5) cells in suspension were sent through a sheath stream where cells were exposed to a laser beam (488nm). EGFP in the cytosol of the cells were excited by the monochromatic blue light and emitted green light at 530nm. The emitted fluorescent light was collected, spectrally filtered and enhanced by a photomultiplier tube in order to provide a value of the green fluorescence emitted (Nolan and Yang, 2007). For a schematic overview of the process, see Figure 2-1.

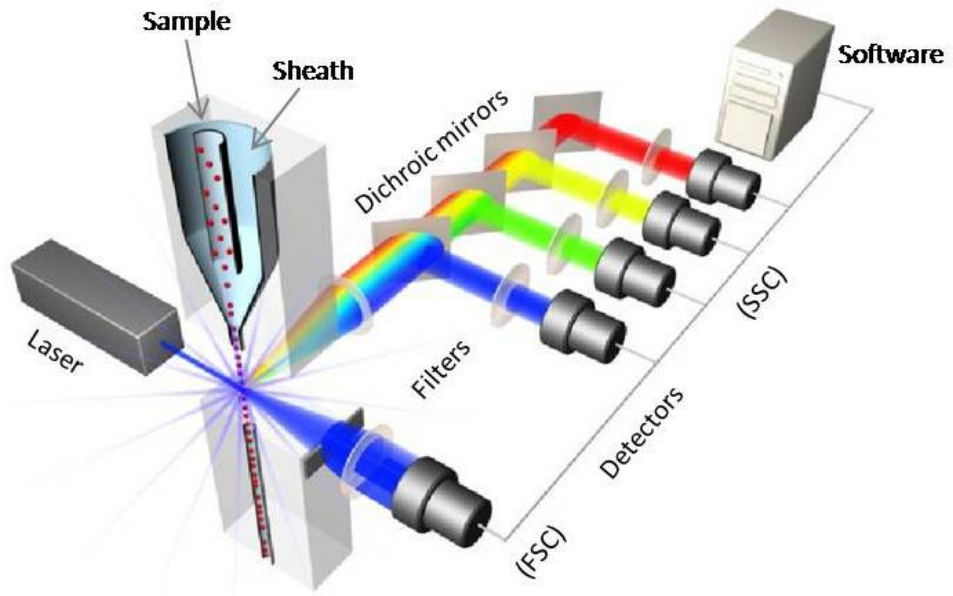


Figure 2-1: Schematic overview of a typical flow cytometer. The cells (Red) in the sample are lined up in the sheath, one cell in diameter, with hydrodynamic focusing. The laser beam (blue), with a specific wavelength, hits one cell at the time and excites the fluorochromes. Scattered and emitted fluorescent light are filtered and collected by specific detectors and the gathered information is sent to a computer and analysed by an appropriate software program.

2.4.1 The Instrument

The flow cytometry experiments were carried out by Hans Christian Dalsbotten Aass at the Flow Cytometry Core Facility (http://www.ulleval.no/flow) at Oslo University Hospital, Ullevål. The cell samples were analysed using a FACS DiVa Vantage Cell Sorter (BD Biosciences, Franklin Lakes, NJ, USA) with a 488nm argon laser supplied with BD FACS DiVa software. For detection of emission through the FITC- (fluorescein isothiocynate) and PI- (propidium iodide) channels, the filters used were 530/30 and 630/22, respectively. A standard number of 10 000 events was recorded for every cell sample.

2.4.2 Parameters and Plots

To understand the plots, some important points need to be explained. First, the pulse registered from each cell as forward scatter (FSC), side scatter (SSC) and/or FITC (green fluorescent light) can be recorded as the pulse height (H), pulse area (A) or pulse Width (W) (see Figure 2-2). Each pulse is converted to a numerical value and plotted in histograms or

two dimensional dot plots (see Figure 2-3: each dot represents a numerical value of the parameters on the two axis). In general, forward scatter (FSC) may reflect the cell size and side scatter (SSC) reflect granularity of the cell interior. Second, the threshold for forward scatter area was set to a pulse value of 5000 to ignore and exclude small particles, small cells and debris that were not interesting in these experiments. An example is indicated by an arrow in Figure 2-3a.

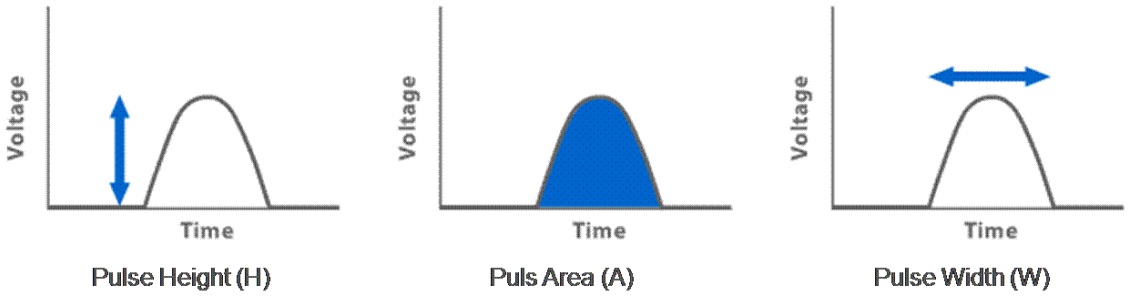


Figure 2-2: Plots visualizing the difference between Pulse- Height, Area and Width registered. Each pulse is converted to a numerical value of one of the three parameters.

Another important point is that most cells have some natural auto fluorescent properties and can be plotted in the same plot as EGFP+ (transfected) cells. The FITC values are plotted against a biexponential (or "Logicle") axis with a linear scale in the lower end and logarithmic scale in the upper end (Herzenberg et al., 2006). Cells recorded with very low or negative values of FITC can now be seen in the same plot as cells with high fluorescent intensity.

2.4.3 Gating approach

Gating is used to select the population of interest and exclude all other particles in the sample. In the beginning of the flow cytometry experiments in this thesis, a specific strategy for gating was used to ensure that the data recorded would in fact come from the cells of interest (HEK293T) and not from dead cells, debris or doublets. The following gating approach was used for all flow experiments carried out in this thesis to ensure that all experiments were done under the same conditions.

2.4.3.1 Determination of viable cells

A transfected control sample (without DNA) of cells was added propidium iodide (PI) to determine dead cells (PI stains the nuclei in dead cells), and to be able to set the gate around the viable cell population. Dead cells that are small and have high granularity, can often be seen as cells with lower forward scatter area (FSC-A) and higher side scatter area (SSC-A) than viable cells. The viable cell population is indicated in blue in the FSC-A /SSC-A plot (Figure 2-3). The cells and debris gated away is indicated in as a purple (Figure 2-3) as a population displaying lower FITC than viable cells and with a partial overlap with the population defined as viable. Excluding dead cells was considered a crucial step of the gating process.

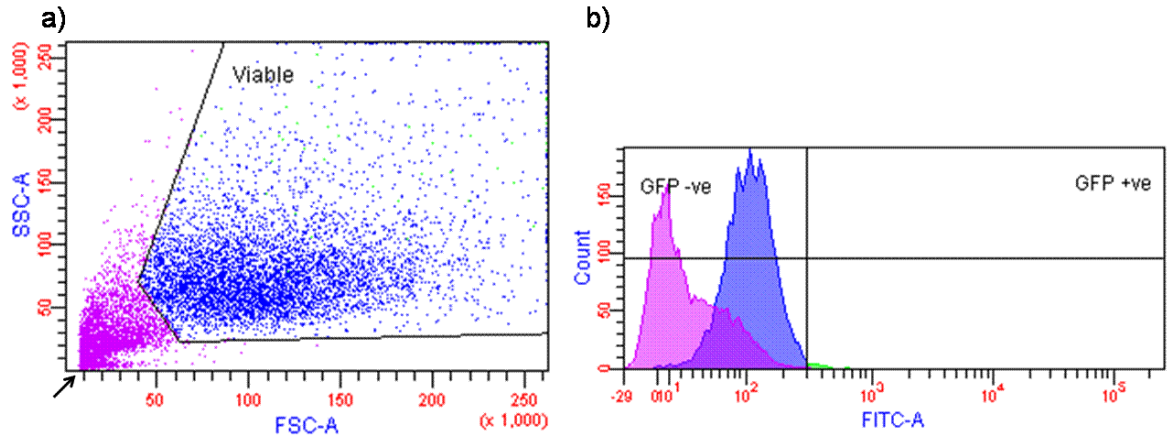


Figure 2-3: Plots showing the gating of viable cells. **a)** The dead cells (purple), identified with PI staining, and other debris, were excluded when the gate was set to the viable population (Blue). The values plotted are side scatter area pulse (SSC-A) and forward scatter area pulse (FSC-A). The arrow indicates the threshold value set for FSC-A (5000) **b)** A histogram showing how dead cells and debris would have interfered with the results as a different population with partial overlap with the viable cells. The values plotted are number of cells (Count) and fluorescent intensity (FITC-A).

2.4.3.2 Discrimination against Doublets

Doublet's (two cells registered as one) can influence the measurements with false strong signals and were therefore gated away in a new plot. The doublets have a pulse area and pulse height equal to the singlets but will give an approximate double pulse width. The doublets were discriminated against, using the SSC-W. Viable singlets are seen inside the gate (blue) in the SSC-A/SSC-W plot (Figure 2-4).

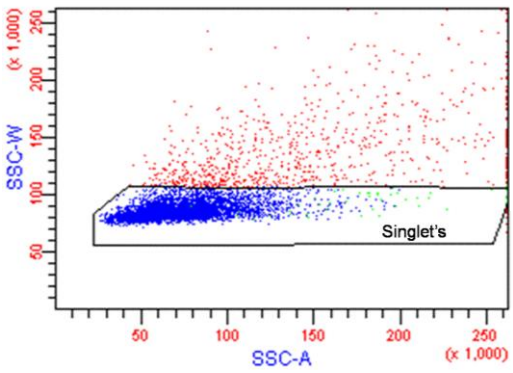


Figure 2-4: Plot showing only viable cells. The doublet's (red) are recognized as cells with higher SSC-W than single cells (Blue). The singlets are gated and selected for further study.

2.4.3.3 Determining EGFP levels in viable single cells.

The control sample of cells transfected without DNA displayed low values of fluorescence. This population was gated and defined as EGFP negative (EGFP-) in the FITC-A/Count plot (Figure 2-5). Cells displaying higher fluorescens were defined as EGFP positive (EGFP+).

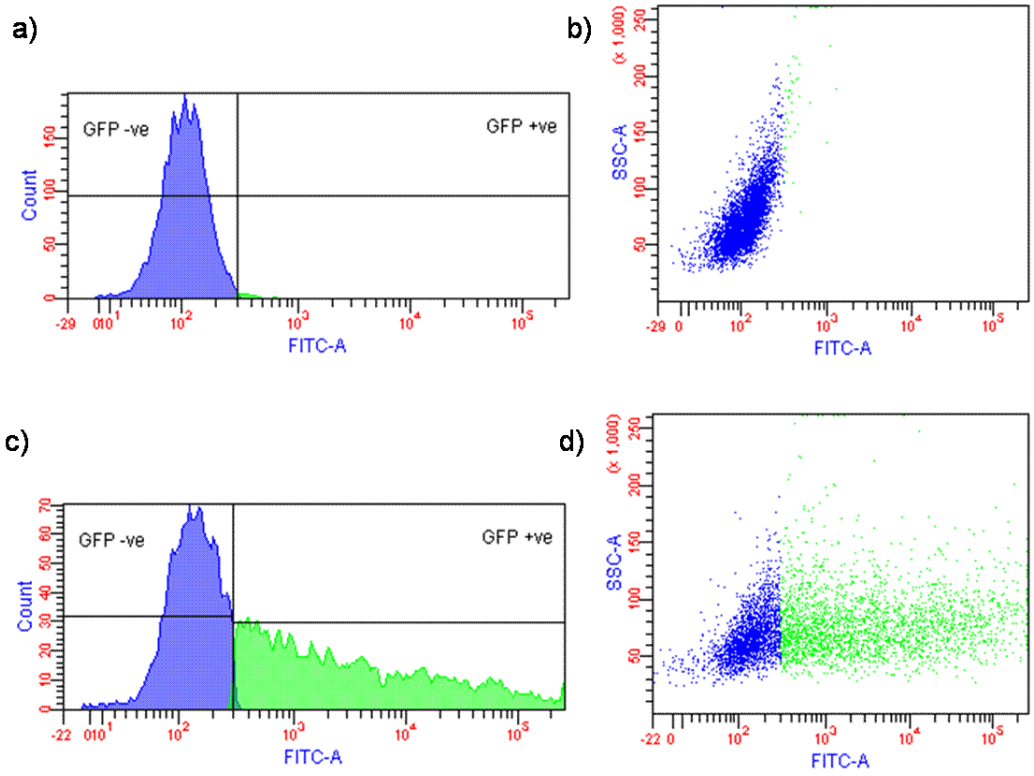


Figure 2-5: Plots showing viable single cells and the results of the gating strategy. **a)** and **b)** showing a representative histogram and a two dimensional dot plot of a typical EGFP- control sample. **b)** and **c)** Showing the same plots from a representative cell sample with transfected with a pmiRPG vector (EGFP+). Blue colour represent cells defined as EGFP- and green colour represent EGFP+ defined cells.

The data plotted graphical in the result chapters is the mean value from the fluorescence intensity recorded from each cell as FITC-A or pulse area as described in section 2.4.2. The mean values were from viable, single EGFP positive cells.

2.5 *In Silico* analysis

Some of the most frequent used programs and software tools are listed and explained below.

2.5.1 DNA manipulation

- National Centre for Biotechnology Information NCBI (http://www.ncbi.nlm.nih.gov/) was used to search for GenBank DNA sequences. The BLAST 2 Sequences program (Blastn) was used to align two DNA sequences.

- Primer 3 (http://frodo.wi.mit.edu/) and Net Primer (http://www.premierbiosoft.com/netprimer/index.html) were used for primer design and quality check, respectively.

- Vector NTI (Invitrogen, Carlsbad, CA, USA) was used as a tool to visualise plasmids with all the functional sequences. The program was also used to search for restriction sites in a vector and to analyse sequencing results using the "Contig Express" tool.

- University of California Santa Cruz, UCSC (http://genome.ucsc.edu/) and Ensemble (http://www.ensembl.org/index.html) were used to search for gene positions in the genome and chromosomes. The Human BLAT Search (UCSC) was used to align known genomic sequences.

2.5.2 miRNA databases

miRNA databases were used to predict possible miRNA matches for the UTR sequences chosen for this study. The criteria for miRNA binding are however not yet fully understood. The various software tools available online use different algorithms (Sethupathy et al., 2006) to predict miRNA target interactions (Table 2-1) and consequently give search results that differ substantially. Five programs were chosen (Table 2-2) to compare and expand searches.

Table 2-1: Summary of the features used by some of the mammalian target prediction programs. Figure modified from Sethupathy *et al*. (2006).

Features	D-microT	miRanda	TargetScanS	PicTar
Sequence				
Perfect seed match rule			×	
Preference for perfect seed match[a]				×
Empirically determined binding rules	×			
Dynamic programming alignment score cutoff		×		
Seed 5′ and/or 3′ flank requirements			×	
Thermodynamics				
ΔG calculations based on traditional RNA folding programs	×	×		
ΔG calculations based on programs for short nucleic acid hybridizations				×
Conservation				
Only between human and rodent species	×	×[b]		
Among human, chimp, rodent, and dog		×[b]	×	×
Residing in an 'island' of conservation			×	

[a]PicTar predict targets with imperfect seed matches, but preferentially predicts targets with perfect seed matches.
[b]miRanda provides the option of running the program under both parameters.

Table 2-2: miRNA target prediction programs used in this thesis.

Program	Website
TargetScanS	http://www.targetscan.org/
miRanda	http://www.microrna.org/microrna/getGeneForm.do
Diana-microT	http://www.diana.pcbi.upenn.edu/cgibin/miRGen/v3/Targets.cgi
PicTar	http://pictar.bio.nyu.edu/cgi-bin/PicTar_vertebrate.cgi
miRBase	http://microrna.sanger.ac.uk/cgi-bin/targets/v4/search.pl

miRNAs predicted to target the selected UTR sequences, were purchased as small miRNA precursors Pre-miR™, which are partially double stranded RNAs that mimic endogenous precursor miRNAs, from Ambion® (Applied Biosystems Inc, Foster City, CA, USA).

3 Results

3.1 Construction of reporter construct pmiRPG

The miRNA binding site reporter vector pmiRPG was constructed (overview see Figure 3-1) from the plasmid pSiRPG (GenBank: DQ465352) an expression vector containing an EGFP encoding gene driven by a phosphoglycerate kinase (PGK) promoter (Størvold et al., 2007). The vector pSiRPG was digested (section 2.2.5), with the restriction enzyme *Eco*RI, ligated (section 2.2.8) and transformed (section 2.1.4) into *E. coli*. The structure of the obtained construct was verified using the restriction enzymes *Eco*RI, *Bam*HI and *Hin*dIII and analyzed on an agarose gel (section 2.2.7) (Mona Lindeberg pers. com.). The resulting vector was named pRPG-1.

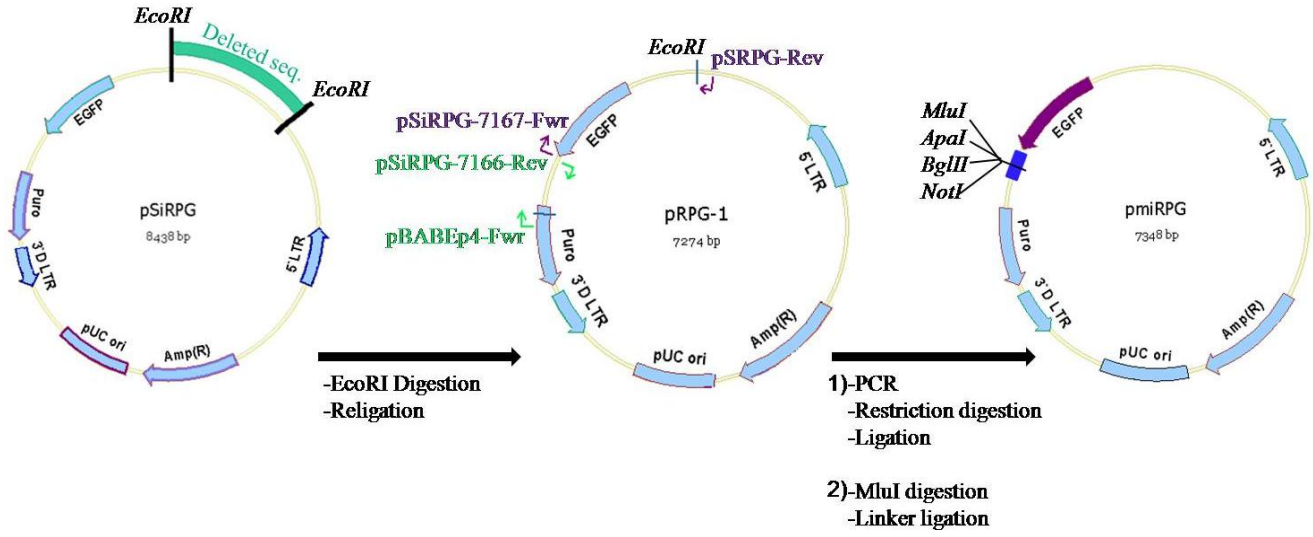


Figure 3-1: Overview of the cloning steps leading to the reporter construct, pmiRPG. See text and Figure 3-2 for detailed descriptions of the cloning steps.

The next steps (Figure 3-2) included several rounds of PCR (section 2.2.10). The purpose was to introduce a 48bp sequence from the 3'UTR of the luciferase reporter gene in the vector pMIR-REPORT™ Luciferase (Applied Biosystems Inc, Foster City, CA, USA). The sequence included the restriction site *Mlu*I that was unique in the new vector. Two primer sets were used in two separate PCR reactions. Two of the primers, *pSiRPG-7167-Fwr* and *pSiRPG-7166-Rev*, were designed to contain the new sequence in tails that were

complementary to each other (indicated in blue in Figure 3-2). The two products, after the first rounds of PCR-reactions (PCR 1 and PCR 2), were mixed together. A second round of PCR was run with *pBABEp4-Fwr* and *pSRPG-Rev* to produce one fused PCR product using the two previous fragments as templates (Figure 3-2). This final PCR product was digested by *Eco*RI and *Bsi*WI. The pRPG-1 vector was digested with the same two enzymes, and the vector fragment was purified from an agarose gel (section 2.2.3). Finally, the two fragments were ligated together and five colonies were selected randomly from the resulting transformation. Purified plasmids were digested in three separate reactions (*Mlu*I, *Hin*dIII and *Eco*RI + *Mlu*I). Analysis on an agarose gel revealed that four out of five clones had the expected fragment sizes (results not shown). Sequencing (section 2.2.11) with primer *pmiRPG-5850-Fwr* and *pmiRPG-6136-Rev* showed that the four clones contained the correct sequence. Clone 2 was chosen for further experiments and was named pRPG-2 (Figure 3-2).

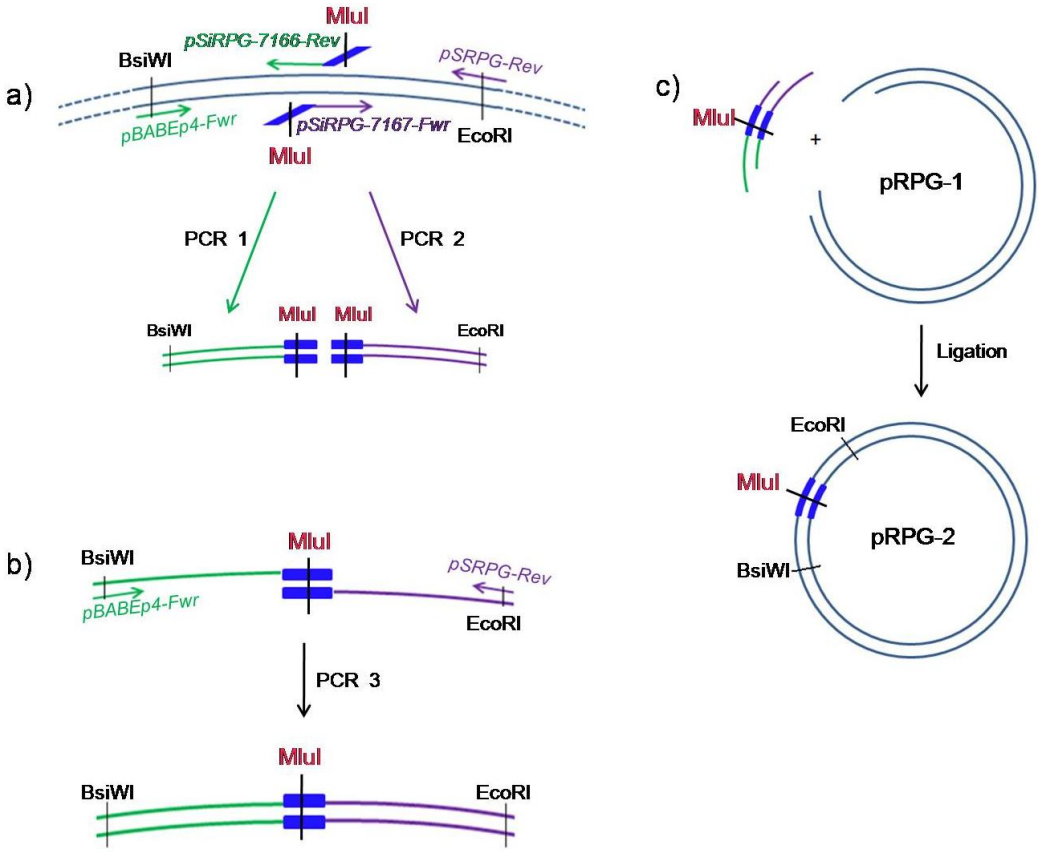


Figure 3-2: Overview of the PCR technique used to construct pRPG-2 by introducing new sequences into pRPG-1: **a)** pRPG-1 was used as a template for two separate PCR reactions (PCR 1 and PCR 2) with primer pairs containing complementary tails (blue) with the new sequence . **b)** The two PCR products were mixed together in a third PCR with the primers *pBABEp4-Fwr* and *pSRPG-Rev*. The two templates annealed to each other in

the complementary tails and the PCR resulted in one fragment **c)** pRPG-1 and the PCR fragment both were digested with *EcoR*I and *BsiW*I and ligated together. The final product was the vector pRPG-2.

The *Mlu*I site in pRPG-2 was used to introduce unique restriction sites to make a functional multiple cloning site in the vector. Two complementary oligo nucleotides containing four restriction sites, *Mlu*I, *Apa*I, *Bgl*II and *Not*I were designed with identical overhangs (Figure 3-3a) for ligation directly into a *Mlu*I digested vector. When this adapter oligo was ligated into the *Mlu*I digested vector, the *Mlu*I site in one end was restored. In the other end the *Mlu*I site was not restored, due to incorrect nucleotides in the start of the site, designed to avoid two *Mlu*I sites in the vector (Figure 3-3).

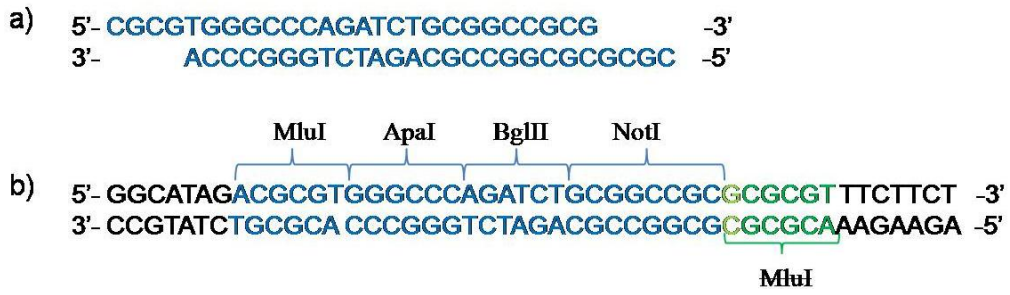


Figure 3-3: **a)** Adapter oligos ligated into vector pRPG-2. The two 5'-ends have identical overhangs with *Mlu*I digested DNA **b)** Sequenced pRPG-2 sequences showing the order of the new restriction sites ligated into the vector.

Ligation reactions were made and transformed into *E. coli*. Nine colonies were randomly selected from the plate. The purified plasmids were digested using *Not*I and *Bgl*II restriction enzymes, both unique in the new multiple cloning site. The results showed that two of the clones (clone 4 and 9) were linearized with both enzymes indicating that the oligo adapter was inserted (Figure 3-4). Sequencing with primers *pmiRPG-5850-Fwr* and *pmiRPG-6136-Rev*, verified that clone 9 had the oligo sequence inserted in the orientation shown in Figure 3-3b. This final reporter construct for miRNA binding site evaluation was named pmiRPG (Figure 3-5).

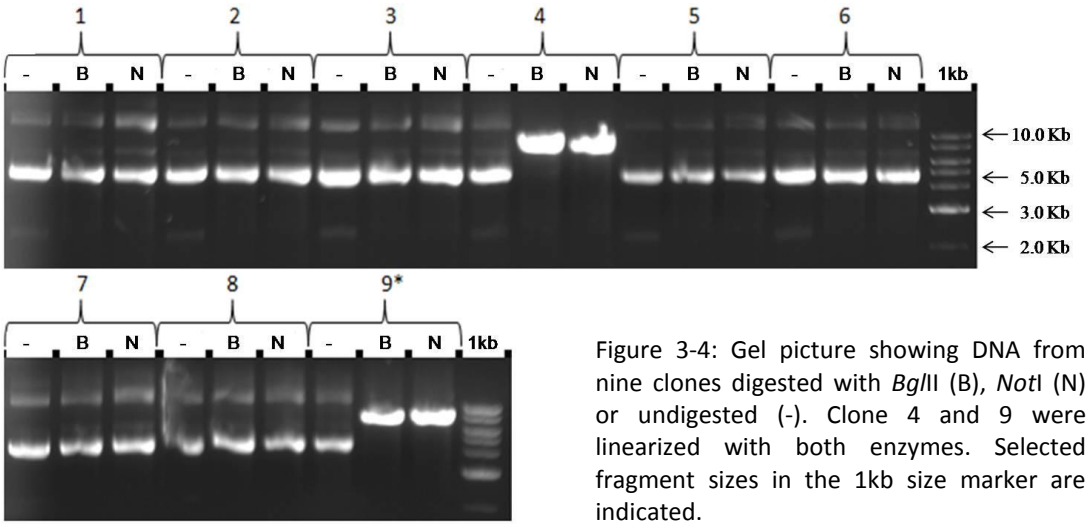


Figure 3-4: Gel picture showing DNA from nine clones digested with *Bgl*II (B), *Not*I (N) or undigested (-). Clone 4 and 9 were linearized with both enzymes. Selected fragment sizes in the 1kb size marker are indicated.

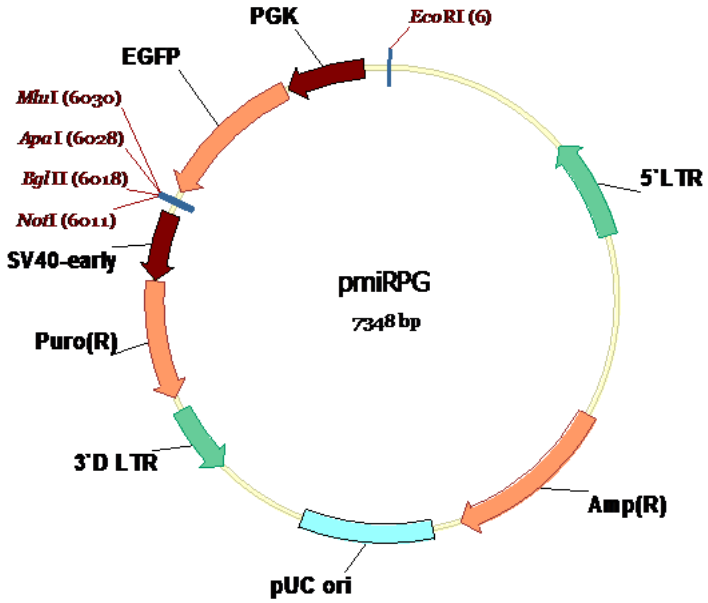


Figure 3-5: The final reporter construct pmiRPG was used in further expression experiment and the 3'UTR sequence of selected genes were cloned into the multiple cloning site using *Mlu*I and *Not*I restriction sites.

3.2 Reporter constructs in pmiRPG

3.2.1 Construction of pmiKRAS-3'U

A construct with the 3'UTR sequence from the *KRAS* gene was constructed in pmiRPG for use in experiments to verify the functionality of the expression vector and for use as a positive control for downregulation of gene expression by miRNAs. Kanamoto et al. (2006) showed that let-7 miRNA anneals to a sequence inside a 1500bp region of the 5kb *KRAS* 3'UTR. This region was cloned in the present work, and the primer sequences used were

modified from Kanamoto et al. (2006) by adding *Mlu*I and *Not*I restriction sites to the forward and the reverse primer, respectively (Appendix, Table 1).These primers (*KRAS-3'UTR-Fwr* and *KRAS-3'UTR-Rev*) were used to amplify the *KRAS* 3'UTR sequence from the cDNA of the breast cell line MCF-7 (provided by Mona M. Lindeberg). The PCR product was analyzed on an agarose gel, and the band of correct size was purified (section 2.2.3). Another round of PCR was performed on the purified product to obtain higher purity and concentration, and the fragments were finally precipitated with ethanol (section 2.2.2).

The *KRAS* 3'UTR fragment was ligated into the pmiRPG vector after the digestion of both with the restriction enzymes *Not*I and *Mlu*I. The pmiRPG vector was digested with two enzymes in two separate tubes. After the first digest, samples from each reaction were analyzed on a 1% agarose gel to verify the linearization (digestion control). The linearized vector fragments were then digested with the second enzyme. The two vector samples digested twice were precipitated with isopropanol to purify the vector.

The *KRAS* 3'UTR fragment and the vector with complementary overhangs were ligated (section 2.2.8) with the T4 ligase (Roche™, Basel, Switzerland). The two vector samples (digested in different order) were ligated with the digested PCR fragment. Ligation controls were set up, one without the PCR fragment and one lacking both fragments and ligase. Additional controls were vectors digested only once, and the *KRAS* fragment digested twice. The ligation reactions were transformed into *E. coli* (Table 3-1).

Table 3-1: Ligation setup and transformation results for the cloning of the selected *KRAS* 3'UTR sequence into reporter construct pmiRPG.

	Ligation:			**Transformation:**		
Plate	**Vector**	**Insert**	**Ligase**	**DNA/50µl bacteria**	**Vol. plated**	**Colonies**
1	pmiRPG/*Not*I/*Mlu*I	*KRAS* 3'UTR/*Not*I/*Mlu*I	T4-ligase	15.0ng	½	96
2	pmiRPG/*Mlu*I/*Not*I	*KRAS* 3'UTR/*Not*I/*Mlu*I	T4-ligase	15.0ng	½	39
3	pmiRPG/*Not*I/*Mlu*I	-	T4-ligase	12.5ng	½	23
4	pmiRPG/*Not*I/*Mlu*I	-	-	12.5ng	½	0
5	pmiRPG/*Not*I	-	T4-ligase	12.5ng	½	57
6	pmiRPG/*Mlu*I	-	T4-ligase	12.5ng	½	250
7	*KRAS* 3'UTR/*Not*I/*Mlu*I	-	T4-ligase	12.5ng	½	0
8	pUC19-control	-	T4-ligase	0.25ng	1/10	13
				Transformation Frequency→ 5 x 10^5		

The two ligations with *KRAS*-3'UTR fragments and ligase, produced 96 and 39 colonies and the reaction with digested vector, with and without ligase produced 23 and 0 colonies. The three last control ligations, *Mlu*I digested and *Not*I digested vector and double digested fragments produced 57, 250 and 0 colonies respectively. Transformed pUC19 produced 13 colonies and thereby reported a poor transformation frequency (5 x 10^5 col/µg DNA). Comparisons between the two first ligations and ligation 3, revealed a slight increase in colonies when the PCR fragments were added to the reactions. Further comparison with the two first ligations and ligation 7 indicated that the double digested vector fragment was sufficiently digested in both ends. The results may indicate that the two first plates could contain positive clones.

Twenty colonies from agar plate 1 and 2 were randomly selected and grown overnight (section2.1.2) and plasmids were purified from the overnight cultures (section2.2.1). An enzyme (*Bst*EII) which cuts twice in the vector and not in the insert sequence was chosen for the first test digestion of the clones. Five clones gave the expected banding pattern after *Bst*EII digestion, showing expected fragments for a vector with the new *KRAS* 3'UTR sequence on the gel (results not shown).

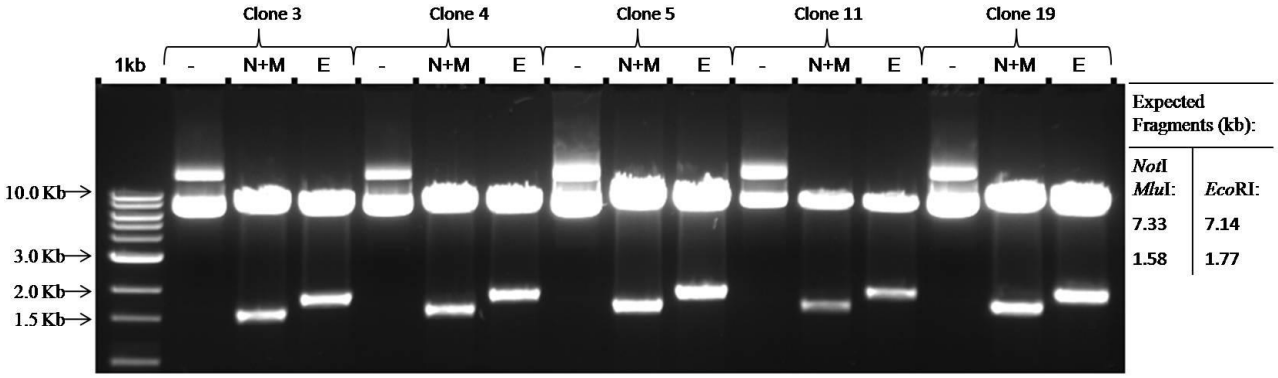


Figure 3-6: Gel picture showing five positive clones undigested (-), and digested with *Not*I (N) + *Mlu*I (M) and *EcoR*I (E). Selected fragment sizes in the 1kb size marker are indicated (left), and the expected fragments from digestion with indicated enzymes are listed (right).

These clones were digested in a second experiment with *Mlu*I and *Not*I in the same reaction, to verify the 1.58kb fragment used, and with *Eco*RI that cuts once in the vector and once in the fragment, as a second control. All five clones showed the expected banding pattern after the two digestion experiments (Figure 3-6) and sequencing with primer *pmiRPG-5850-Fwr*

and *pmiRPG-6136-Rev* revealed that clone five contained the correct fragment (Figure 3-7). This clone was named pmiKRAS-3'U.

Figure 3-7: Clone 5 sequence (blue) sequenced with indicated primers and aligned with the *KRAS* (Black) gene sequence (GenBank: M54968). The KRAS-3'UTR sequence flank the enzyme sites (Red) in both sequences read.

3.2.2 Construction of pmiCXXC4-3'U and pmiWDR79-3'U

Reporter construct pmiCXXC4-3'U and pmiWDR79-3'U were constructed by insertion of the 3'UTR sequences from the *CXXC4* and the *WDR79* gene into pmiRPG. The 136bp 3'UTR sequence from *CXXC4* was amplified (section 2.2.10) from Human (Male) Genomic DNA (Promega, Madison, WI, USA) using a primer pair, *CXXC4-3'UTR-Fwr* and *CXXC4-3'UTR-Rev* (Appendix, Table 1), designed to flank the entire 3'UTR sequence and containing restriction sites for *Mlu*I and *Not*I in the tails. The short 3'UTR sequence from *WDR79* was purchased as complementary oligo nucleotide sequences designed to include five base pairs upstream and downstream of the 16bp UTR sequence, and with overhangs in the ends compatible with *Mlu*I and *Not*I restriction sites (Appendix, Table 2).

The pmiRPG vector and the amplified UTR fragment from *CXXC4* were digested with *Mlu*I and *Not*I and precipitated (section 2.2.2). The digested vector and the two UTR sequences were ligated together in separate reactions (section 2.2.8). Three control ligation reactions were made, one without the PCR fragments, one without ligase or fragments, and one with vectors digested only once (Table 3-2).The ligation reactions were transformed into *E. coli*.

Table 3-2: Ligation setup and transformation results for the cloning of *CXXC4* 3'UTR and *WDR79* 3'UTR into reporter construct pmiRPG.

	Ligation:			**Transformation:**		
Plate	**Vector**	**Insert**	**Ligase**	**DNA/50µl bacteria**	**Vol. plated**	**Colonies**
1	pmiRPG/*Mlu*I/*Not*I	*CXXC4* 3'UTR/*Not*I/*Mlu*I	T4-ligase	26.9ng	½	18
2	pmiRPG/*Mlu*I/*Not*I	*WDR79* Oligo seq. x 2	T4-ligase	25.8ng	½	27
3	pmiRPG/*Mlu*I/*Not*I	-	T4-ligase	25.0ng	½	4
4	pmiRPG/*Mlu*I/*Not*I	-	-	25.0ng	½	0
5	pmiRPG/*Mlu*I	-	T4-ligase	25.0ng	½	93
6	pUC19-control	-	T4-ligase	0.25ng	1/10	5
				Transformation Frequency → 4 x 10^4		

The two ligations with UTR fragments and ligase produced 18 and 27 colonies and the reaction with digested vector, with and without ligase produced 4 and 0 colonies. The vector only digested once by *Mlu*I produced 93 colonies and the transformed pUC19 control produced 5 colonies and thereby reported a poor transformation frequency (4 x 10^4). Comparisons between the two first ligations and ligation 3, revealed an increase in colonies when the PCR fragments were added to the reactions. Comparisons between vector digested once and twice (ligation 3 and 5) indicated that the vector was sufficiently digested with both enzymes. The results indicated that the resulting agar plates from ligation with *CXXC4* fragment and *WDR79* fragment could contain positive clones.

Plasmid DNA from ten of the 18 colonies on agar plate number one was digested with restriction enzyme *Hind*III which cuts twice in the vector sequence. Four clones (clone 5, 8, 9 and 10) gave the expected banding patter with *Hind*III and were digested in a second reaction with restriction enzymes *Not*I and *Mlu*I to visualise the fragment of correct size on the gel. Three of the four clones showed a fragment with a size corresponding to the 0,17kb expected from the *CXXC4* 3'UTR (Figure 3-8a).

Six colonies were selected randomly from agar plate number two (Table 3-2). Plasmids were purified (section 2.2.1) and digested with restriction enzymes *Not*I and *Mlu*I. 1.5µg of each clone were digested to obtain an amount of the oligo nucleotide sequence visible on a 2% agarose gel (approximately 7.5ng). The six clones showed a fragment with a size corresponding to the 0.037kb expected from *WDR79* 3'UTR (Figure 3-8).

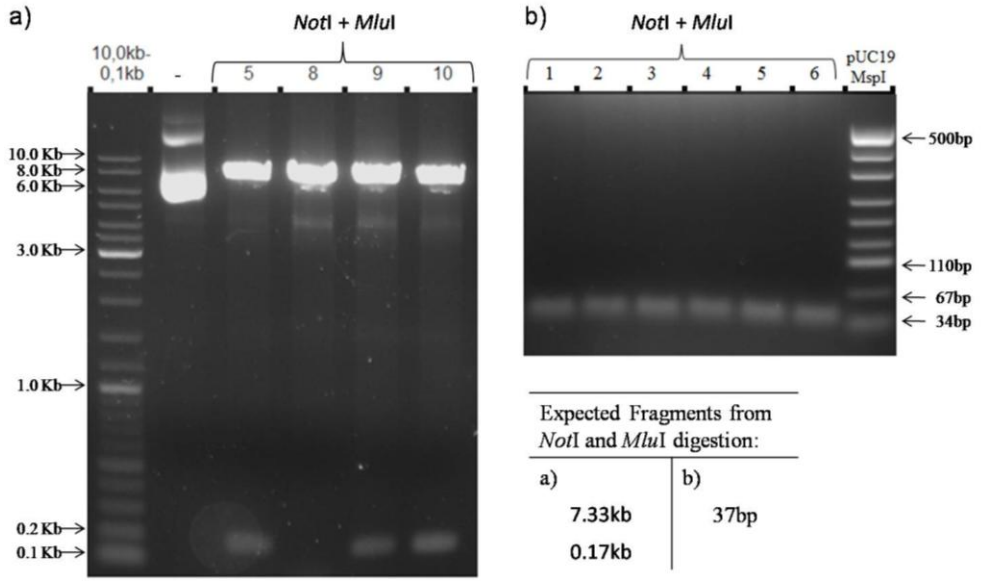


Figure 3-8: Gel pictures from digestion experiments with potential positive CXXC4 3'UTR clones (a) and WDR79 3'UTR clones (b). **a)** 1.5% Agarose gel showing the results of an undigested (-) control and clones digestion with *Not*I and *Mlu*I enzymes. Selected fragment sizes in the 10-0.1kb size marker are indicated. **b)** 2% Agarose gel showing the result of digestion with *Not*I and *Mlu*I enzymes. Selected fragment sizes in the pUC19/*Msp*I size marker are indicated. The **table** shows the expected fragments for positive clones in the two experiments.

Clone 5 from ligation 1, and clone 1 from ligation 2 (Table 3-2 and Figure 3-8) was sequenced (section 2.2.11) with primer *pmiRPG-5850-Fwr* and *pmiRPG-6136-Rev* to verifying that the correct UTR sequences from *CXXC4* and *WDR79* were inserted in the vectors, and the constructs were named pmiCXXC4-3'U and pmiWDR79-3'U (Appendix, Figure 2 and 3).

3.2.3 Construction of pmiWWOX-3'U

The construction of reporter construct pmiWWOX-3'U from pmiRPG was carried out by inserting the 870bp 3'UTR sequence from the gene *WWOX* into the *Mlu*I and *Not*I sites in pmiRPG. The construct from this cloning was sequenced with primer *pmiRPG-5850-Fwr* and *pmiRPG-6136-Rev* to confirm the presence of the *WWOX* 3'UTR sequence (Appendix, Figure 4). The new clone was named pmiWWOX-3'U (Mona M. Lindeberg pers. com.).

3.3 Results from Flow Cytometry experiments

The following sections presents results from flow cytometry analysis performed on transiently transfected HEK293T cells (section 2.3.4), using vectors expressing EGFP. All experiments were performed in duplicates, and the cells were analysed using flow cytometry as described in section 2.4.The gating strategy is described in section 2.4.3.

3.3.1 EGFP expression from vectors

The first experiment was carried out to ensure that the expression of EGFP in the vectors was not altered during the cloning steps. The cells were transfected with the three vectors, pSiRPG, pmiRPG and pmiWWOX-3'U. The cells were analysed using flow cytometry after 24 hours. The results obtained are presented in Figure 3-9.

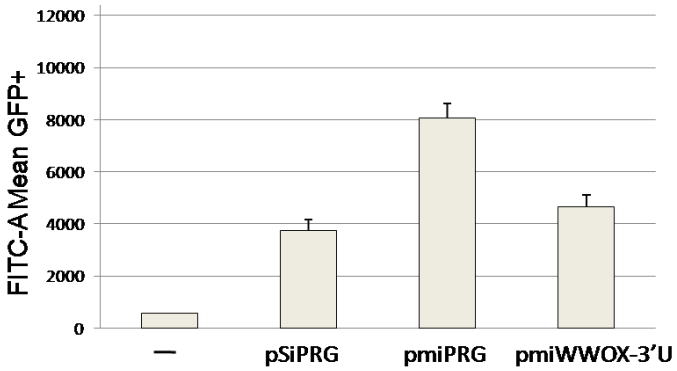


Figure 3-9: Plot showing fluorescent intensity (FITC) from cells transfected with pSiRPG, pmiRPG, pmi*WWOX*-3'U and no DNA (-) as indicated. The data plotted are Mean intensities (FITC-A) from cells defined as EGFP+.

The results show the three vectors pSiRPG, pmiRPG and pmiWWOX-3'U with EGFP expression (FITC) significantly above the normal auto fluorescence from the control sample (cells transfected without DNA). Although there were differences in the fluorescence intensity between the three vectors, showing pmiRPG with an average intensity of about the double of the intensities recorded from pSiRPG and pmiWWOX-3'U, the expression functions of the vectors were shown to remain intact through the cloning steps and these constructs were used for further experiments.

3.3.2 Investigation of EGFP expression from pmiKRAS-3'U and regulation by Let-7a

Two parallel experiments were carried out, to verify the EGFP expression from reporter construct pmiKRAS-3'U, and to investigate the effect of miRNA let-7a on the expression of the vector. The cells were transfected with pmiKRAS-3'U only, and pmiKRAS-3'U cotransfected with 100pmol of siGFP Duplex (Thermo Scientific, Wilmington, De, USA) targeting the EGFP transcript (Appendix, Table 2) and 100pmol Let-7a miRNA. A cell sample transfected without DNA was included as a control. The cells were analysed using flow cytometry, after 24 hours (Figure 3-10).

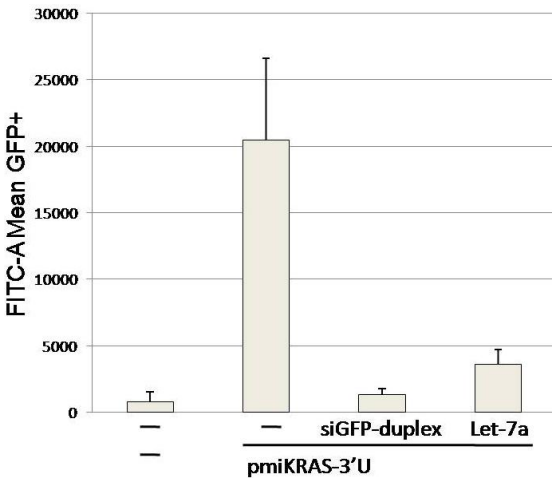


Figure 3-10: Plot showing fluorescent intensity (FITC) from cells transfected with pmi*KRAS*-3'U or no DNA (-) and cells cotransfected with pmi*KRAS*-3'U and SiGFP-duplex or Let-7a miRNA. The data plotted are Mean intensities from cells defined as EGFP+.

The flow cytometry results showed significant fluorescence from HEK293T cells transfected with the vector pmiKRAS-3'U. The EGFP expression from cells cotransfected with pmiKRAS-3'U and SiGFP-duplex was significantly down regulated, compared to the control cells transfected with pmiKRAS-3'U (Figure 3-10). This result verified that siGFP-duplex targets the *EGFP* gene in the vector. Thus, the siGFP-duplex can be used in further experiments as a control for down regulation of the EGFP expression. The reduced EGFP expression in cells transfected with let-7a indicates down regulation of *KRAS* by Let-7a, as previous shown by others (Johnson et al., 2005).

3.3.3 *In Silico* miRNA target prediction

Five online miRNA target prediction programs (section 2.5.2) were used in combination to predict miRNAs that target the 3'UTR sequences of the transcripts *WWOX, CXXC4, and WDR79*. The gene symbol/names were plotted into the programs and the human genome was selected for search of miRNAs. The programs showed a considerable variation in the miRNAs that they predicted and a decision was made to choose the miRNAs that were predicted by two or more programs. The miRNAs chosen for experiments in this thesis are listed in Table 3-3.

Table 3-3: Overview of miRNAs selected for further study in this thesis. The miRNAs were predicted by two or more of the indicated programs, or selected for the reported conservation. Only one program predicted miRNAs for *WDR79*, and these miRNAs were both selected.

Target	Prog.	miRNA Stem-loop	miRNA Mature	miRNA-Mature Sequence 5'-3'
KRAS	-	hsa-let-7a-1	hsa-let-7a	ugagguaguagguuguauaguu
WWOX	1,2	hsa-mir-758	hsa-miR-758	uuugugaccugguccacuaacc
	1	hsa-mir-516b-1	hsa-miR-516b*	ugcuuccuuucagagggu
	1,2	hsa-mir-134	hsa-miR-134	ugugacugguugaccagagggg
CXXC4	1,2	hsa-mir-205	hsa-miR-205	uccuucauuccaccggagucug
	1,2	hsa-mir-570	hsa-miR-570	cgaaaacagcaauuaccuuugc
	1,2,4	hsa-mir-105-2	hsa-miR-105	ucaaaugcucagacuccuguggu
	1,5	hsa-mir-532	hsa-miR-532-5p	caugccuugaguguaggaccgu
WDR79	1	hsa-mir-487a	hsa-miR-487a	aaucauacagggacauccaguu
	1	hsa-mir-548d-2	hsa-miR-548d-3p	caaaaaccacaguuucuuuugc

The miRNAs listed are predicted by: TargetScanS[1], miRanda[2], Diana-microT[3], PicTar[4] and miRBase[5]

3.3.4 Co-transfection of reporter constructs with miRNAs

miRNA binding sites were predicted in the 3'UTR of *WWOX* for the miRNAs hsa-miR-516b*, hsa-miR-134 and hsa-miR-758 (Table 3-3). These miRNAs and let-7a, were used in an experiment, cotransfected with pmiWWOX-3'U into HEK293T cells. 100pmol of each miRNA were used in the experiment.

Flow analysis 24 hours after transfection, revealed down regulation of the EGFP expression (FITC) in cells transfected with pmiWWOX-3'UTR and the siGFP-duplex, compared to the control cells transfected with the vector only. The results also showed EGFP down regulation in cells transfected with the three predicted miRNAs and let-7a (Figure 3-11).

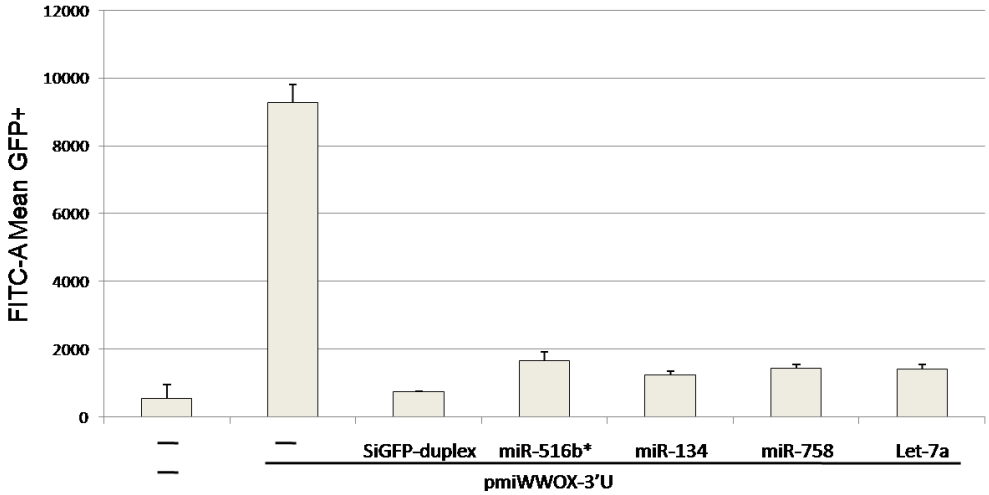


Figure 3-11: Plot showing fluorescence intensity (FITC) from cells transfected with pmiWWOX-3'U or no DNA, and cells cotransfected with pmiWWOX-3'U and the four indicated miRNAs. The data plotted are Mean intensities from cells defined as EGFP+.

This experiment needed an additional experiment to confirm that the miRNAs tested specifically downregulated *WWOX 3*'UTR. The next experiment was designed with a wider selection of miRNAs and the miRNAs were added to the cell sample in two different amounts. One purpose of the new experiment was to investigate whether the reduction of the EGFP-signal from the cells cotransfected with miRNA was a result of specific down regulation, or a result of unspecific down regulation due to the concentration of miRNA used in the experiment (100pmol).

The six miRNAs hsa-miR-516b*, hsa-miR-134, hsa-miR-758, hsa-miR-205, hsa-miR-578 and hsa-miR-487a were cotransfected with pmiWWOX-3'U into HEK293T. Two amounts of each miRNA were used, 100pmol and 25pmol. The cells were analysed by flow cytometry after 24 hours (Figure 3-12).

The result showed that all six miRNAs down regulated the expression of EGFP with 100pmol of added miRNAs. When the amount of miRNAs added was decreased to 25pmol, the repression decreased for five miRNAs while the last miRNA, miR-487a showed the repression as when 100pmol miRNA was used.

In summary, the information from these experiments may indicate that unspecific regulation by miRNAs could be diminished with titration experiments to optimise the concentration of miRNAs added. The results from the samples with miR-487a could indicate that this miRNA

regulate the *WWOX* gene, and further that the three miRNAs predicted to regulate the *WWOX* gene, were false predictions.

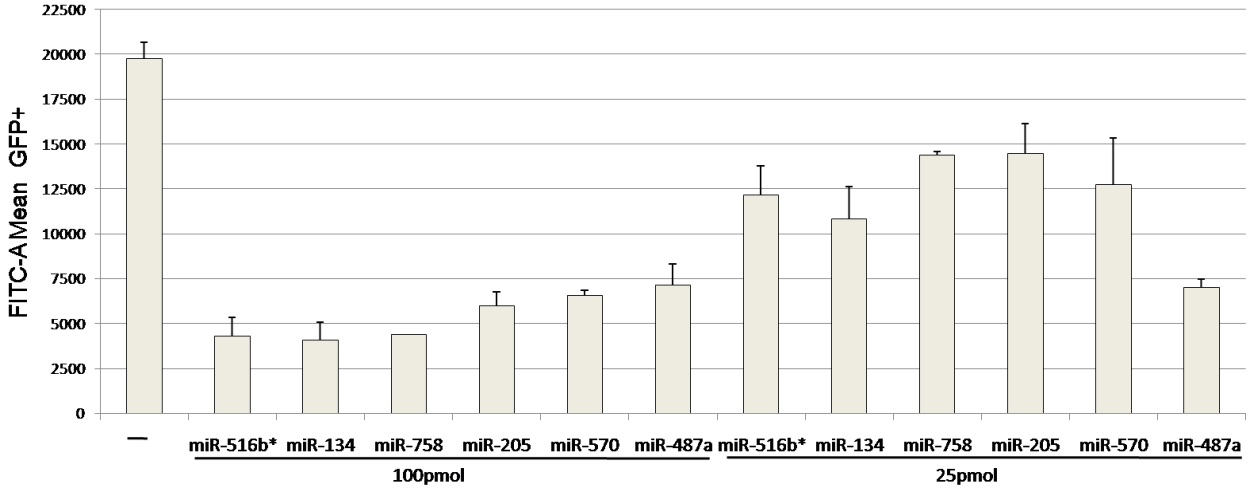


Figure 3-12: Plot showing fluorescence intensity (FITC) from cells transfected with pmiWWOX-3'U and cells cotransfected with pmi*WWOX*-3'U and indicated miRNAs in two amounts (100pmol and 25pmol). The data plotted are Mean intensities from cells defined as EGFP+.

The experiment with the pmiWWOX-3'UTR had raised important questions concerning the design of the experiments and in addition, there were still two expression vectors that remained unexplored in transfection experiments, the pmiCXXC4-3'U and pmiWDR79-3'U. The importance of exploring the fine tuning of the experiments led to the decision to leave the unexplored vectors and focus on the follow up of experiments already done with the pmiWWOX-3'U vector and continue with this vector to optimize the experiments.

3.3.5 Titration experiments may increase specificity

Two separate titration experiments were carried out to optimize the miRNA amounts added in experiments where miRNAs were cotransfected with the reporter vector. The first experiment was done with pmiWWOX-3'U cotransfected with two fold dilutions of let-7a starting at 100pmol per well. The cells were analysed with flow cytometry after 24 hours (Figure 3-13).The result showed an increase in EGFP expression with decreased let-7a concentration from 50pmol to 12.5pmol added.

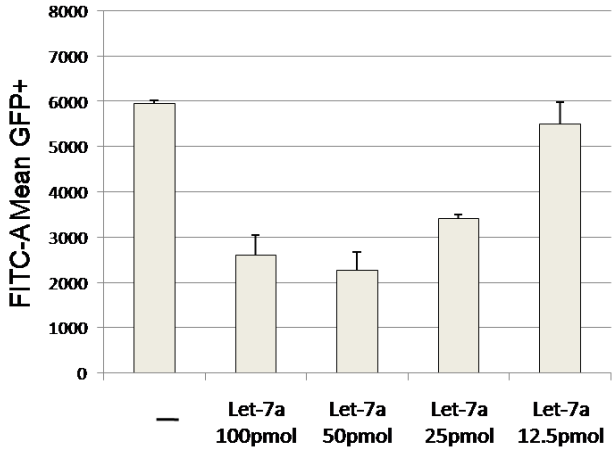


Figure 3-13: Plot showing fluorescent intensity (FITC) from cells transfected with pmi*WWOX*-3U and Let-7a in the amounts indicated. The data plotted are Mean intensities from cells defined as EGFP+.

Two parallel experiments were designed to explore the consequence of even lower amounts of Let-7a cotransfected with pmiWWOX-3'U. The experiments carried out similar to the previous experiment, but with four-fold dilutions of let-7 starting at 100pmol and ending at 1.56pmol, gave the results shown in Figure 3-14.

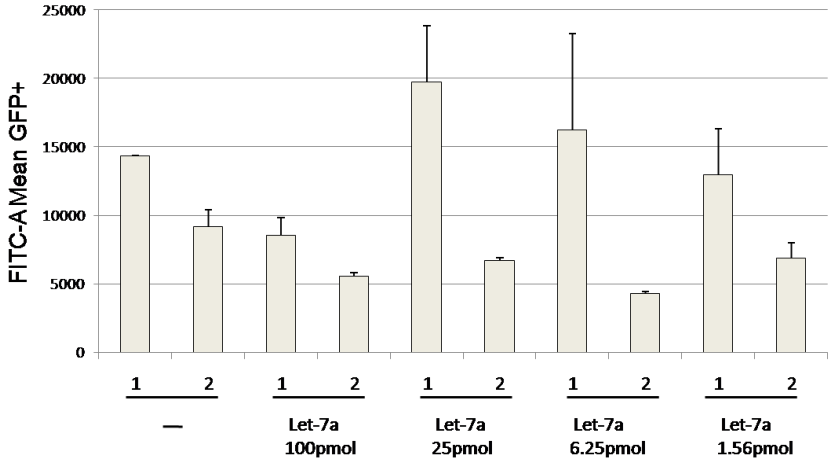


Figure 3-14: Plot showing fluorescent intensity (FITC) from cells in two separate experiments (experiment 1 and 2), transfected with pmi*WWOX*-3U and indicated amounts of Let-7a. The data plotted are Mean intensities from cells defined as EGFP+.

The results of the two experiments showed that all four amounts of let-7a lead to decreased EGFP signals in experiment 2, but variable results were obtained in experiment 1. Therefore, the optimal miRNA concentration where the miRNAs specifically targeting a transcript decrease the expression, and miRNAs that do not target the transcript give limited effect on the expression, was not found. The titration experiments may indicate that the method used to measure the EGFP expression in the experiments, should be investigated in order to obtain less variation between experiments. One possibility may be to increase the EGFP expression in the system, facilitating single cell expression measurements instead of the population based measurements used in this thesis (section 2.4.3.3).

3.3.6 Comparison of promoter PGK and CMV in transient transfection

The vector pEGFP-N3 (GenBank: U57609), constructed with an *EGFP* gene expressed from the CMV (Human Cytomegalovirus) immediate early promoter (Muller et al., 1990), was chosen to compare the expression of *EGFP* from this strong promoter to *EGFP* expressed from the PGK promoter in pmiRPG. The *EGFP* gene sequences were identical in the two vector systems.

HEK293T cells were transfected with pSiRPG, pmiRPG, pmiWWOX-3'U and pEGFP-N3. The results from flow analysis after 24 hours (Figure 3-15) showed that the average fluorescence intensity in the cell population transfected with pEGFP-N3 was recorded to be about eight times higher than the intensity of the cells transfected with vectors with a PGK promoter.

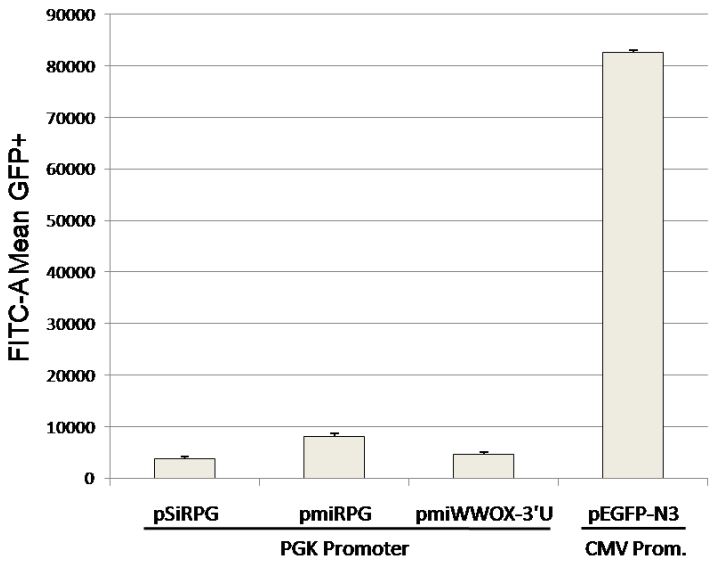
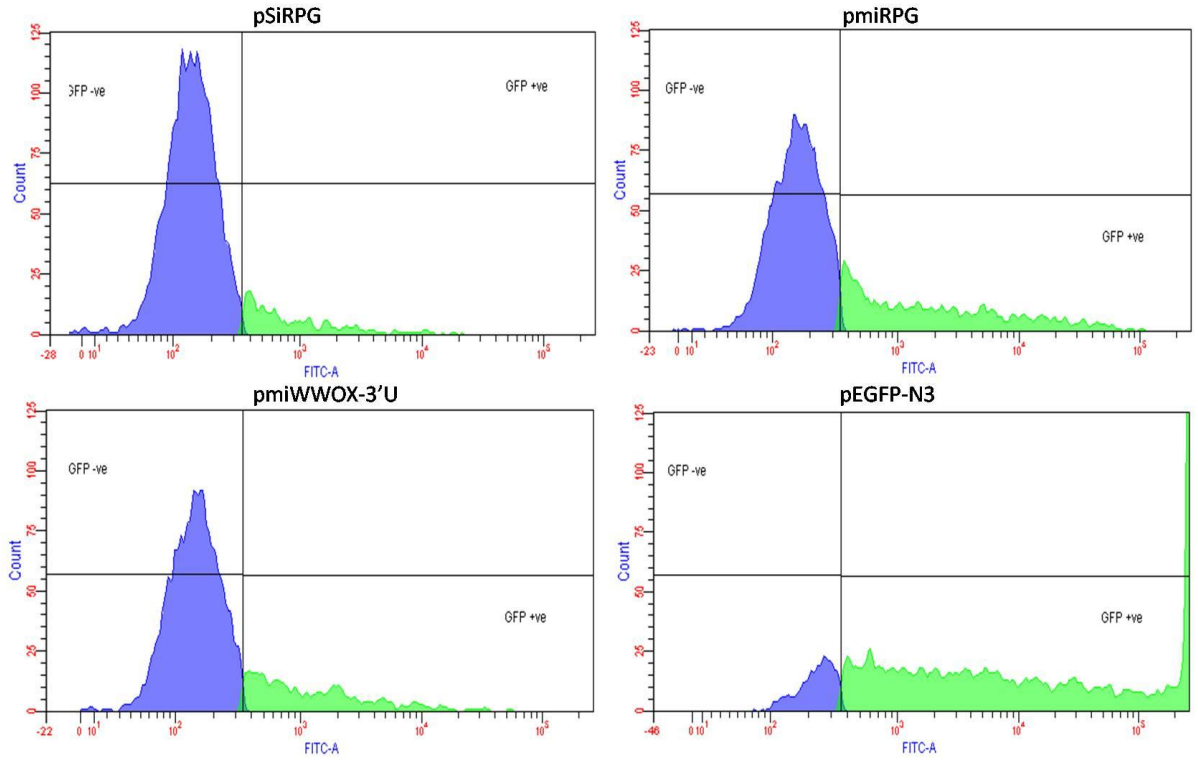


Figure 3-15: Plot showing fluorescence intensity (FITC) from cells transfected with pSiRPG, pmiRPG and pmi*WWOX*-3'U containing the PGK promoter, and pEGFP-N3 containing the CMV promoter. The data plotted are Mean intensities from cells defined as EGFP+.

The difference in EGFP expression between the vector with PGK promoter compared to the pEGFP-N3 could be seen directly in the histograms from the flow experiments (representative examples shown in Figure 3-16), where the green part of the histogram

indicate the EGFP positive cells, and blue indicate EGFP negative cells.



	pSiRPG		pmiRPG		pmiWWOX-3'U		pEGFP-N3	
% EGFP-	89.0	88.7	72.0	71.7	80.0	80.6	15.7	15.5
% EGFP+	11.0	11.3	28.0	28.3	20.0	19.4	84.3	84.5

Figure 3-16: Representative histograms showing the results from flow analysis of HEK293T cells transfected with three vectors containing a PGK promoter (pSiRPG, pmiRPG and pmiWWOX-3'U), and one vector with a CMV vector (pEGFP-N3). Blue colour represent cells defined as EGFP negative and green colour represent EGFP positive cells. The fluorescence intensity (FITC-A) are plotted on a biexponential axis. The **table** shows the percentage distribution of cells in the EGFP positive or EGFP negative population for duplicate samples.

The three histograms from cells transfected with vectors pSiRPG, pmiRPG and pmiWWOX-3'U had the majority of the cells in the population defined as EGFP-negative and assumed untransfected. In the histogram from the cells transfected with pEGFP-N3, the vector with CMV driven EGFP, the majority of the cells was shifted to the area defined as EGFP+, with a peak confined to the highest fluorescence intensity.

The reason why the distribution of cells in the histograms differs substantially between the two vector systems could be that pSiRPG, pmiRPG and pmiWWOX-3'U must be transfected into each cell in higher quantities than pEGFP-N3 to make the cells display a fluorescence intensity above the normal auto fluorescence and to define the cells as EGFP positive.

pEGFP-N3 might give the cells fluorescence properties even in only a few copies and thereby seems to have a higher transfection efficiency compared to the three other vectors in this experiment.

3.3.7 Analysis 24 and 48 hours after transfection

The fluorescence intensity from expressed EGFP coded by reporter construct pSiRPG, pmiRPG, pmiWWOX-3'U and pEGFP-N3 were investigated in an experiment where transfected cells were analysed 24 and 48 hours after transfection. Cells were plated at day one, and new cells were plated out on day two. The cells plated out on day one and two were transfected with the vectors on day two and three, respectively. The cells were analysed at the same time, day four, under the same conditions using flow cytometry, and the results are presented in Figure 3-17.

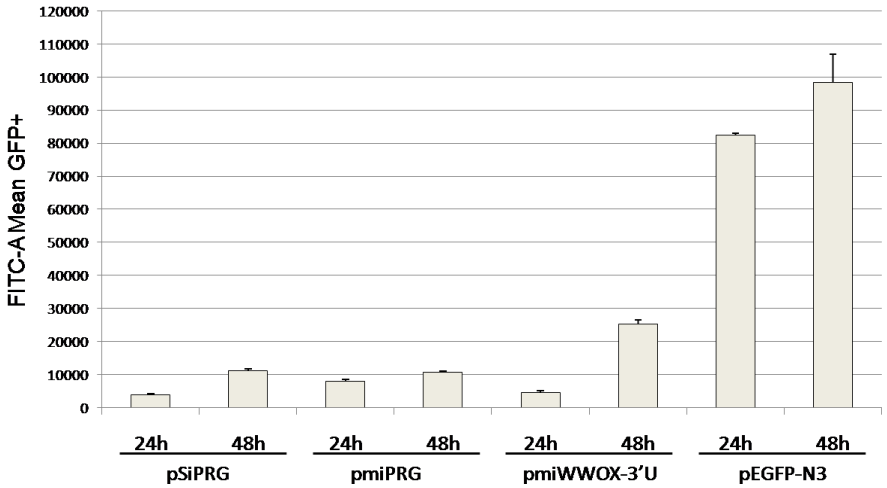


Figure 3-17: Plot showing fluorescence intensity (FITC) from cells transfected with pSiRPG, pmiRPG, pmi*WWOX*-3'U and pEGFP-N3 and analyzed after 24 and 48 hours. The data plotted are Mean intensities from cells defined as EGFP+.

The four vectors showed increased fluorescence intensities after 48 hours compared to 24 hours. The expression from constructs containing the PGK promoter however, was still low compared to the pEGFP-N3 vector with CMV promoter after both 24 and 48 hours. The experiment showed that prolonged incubation time (from 24 to 48 hours) for cells transfected with expression vectors, did result in a modest increase of EGFP expression, but probably not enough to facilitate single cell expression measurements. The results may suggest that an exchange from PGK promoter to a CMV promoter could give better expression of the *EGFP* gene in the miRNA binding site reporter vectors.

3.3.8 Investigating constructs with PGK and CMV driven EGFP-reporters

Differences between the PGK- and the CMV promoter driven EGFP-gene were further investigated in an experiment where the three vectors pmiRPG, pmiWWOX-3'U and pEGFP-N3 were cotransfected with siGFP duplex, hsa-miR-134 and the siGENOME Non-Targeting siRNA#1 (Appendix, Table 3) that do not target any genes in the genome. The Non-Targeting siRNA#1 contain no significant sequence similarity to the *EGFP* gene.

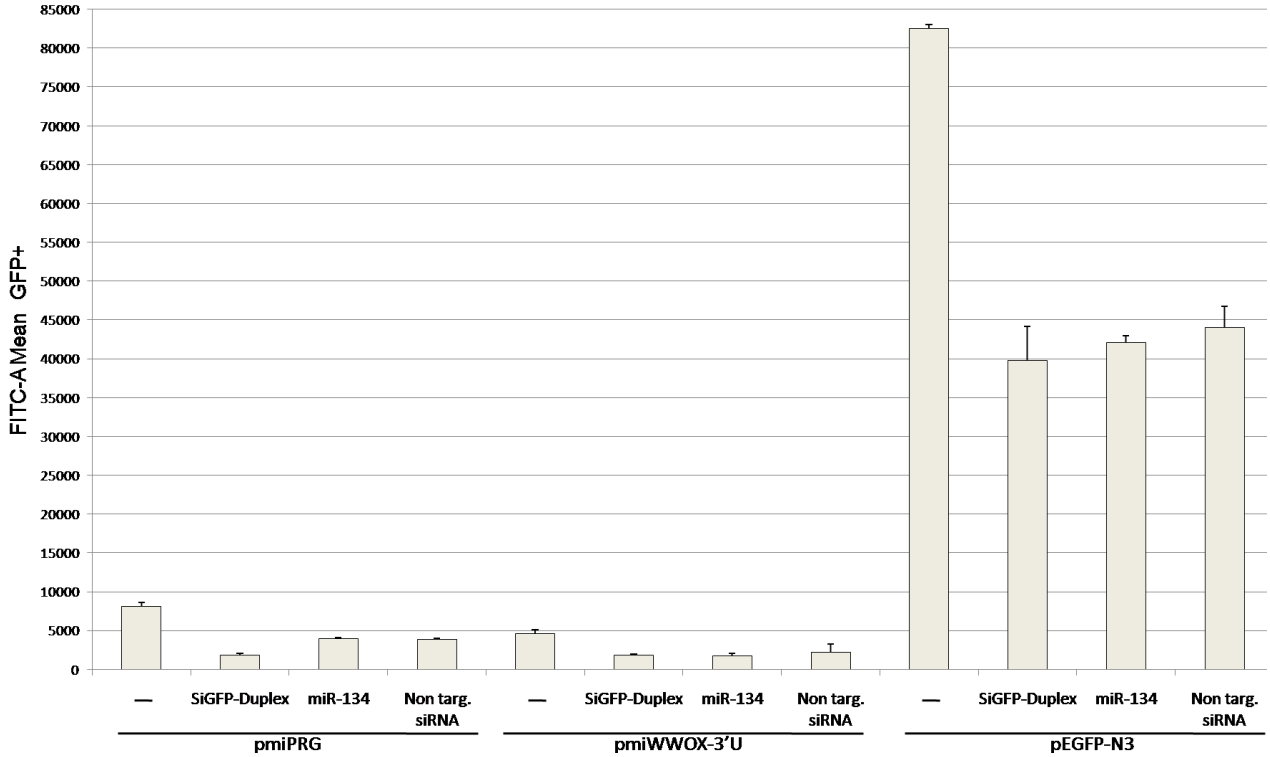
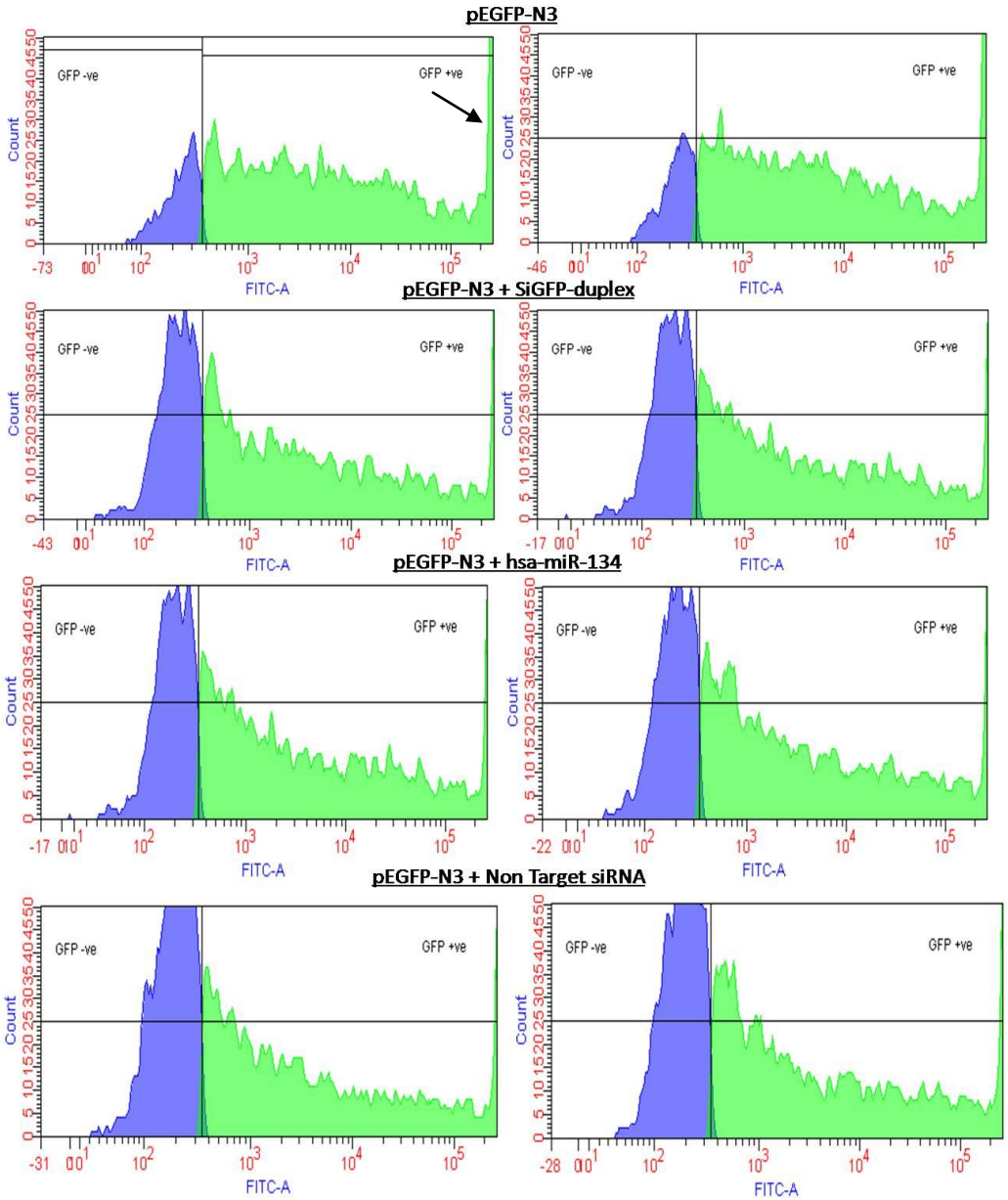


Figure 3-18: Plot showing fluorescence intensity (FITC) from cells transfected with pmiRPG, pmiWWOX-3'U and pEGFP-N3. The three vectors are also cotransfected with SiGFP-duplex, hsa-miR-134 and a non targeting siRNA. The data plotted are Mean intensities from cells defined as EGFP+.

Again the results (Figure 3-18) showed a considerable higher EGFP expression from vectors containing the CMV promoter compared to the PGK promoter. The siGFP-duplex reduced the EGFP expression in cells transfected with all three vectors. The hsa-miR-134 and the non targeting siRNA reduced the signals from all three vectors significantly, which may indicate that further optimalization of miRNA/siRNA concentration is needed.



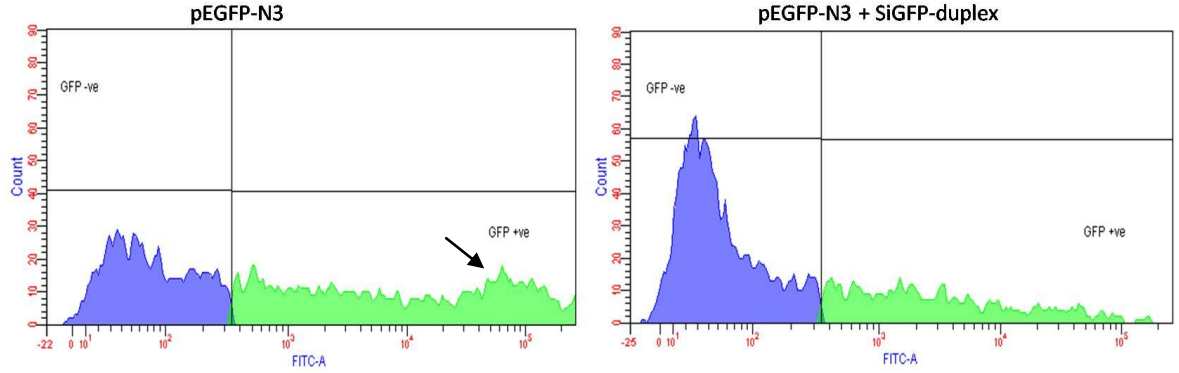
	-		siGFP-duplex		miR-134		Non-target siRNA	
% EGFP-	15.7	15.5	34.1	40.0	36.8	35.5	43.7	41.2
% EGFP+	83.3	84.5	65.9	60.0	63.2	64.5	56.3	58.8

Figure 3-19: histograms showing the results from flow analysis of transfected HEK293T cells. The cells are transfected with pEGFP-N3 and cotransfected with the vector and siGFP-duplex, miR-134 miRNA and a non target siRNA. Blue colour represent cells defined as EGFP- and green colour represent EGFP+ defined cells. The fluorescence intensity (FITC-A) are plotted on a biexponential axis.

Experiments with the pEGFP-N3 vector revealed that in contrast to the experiments carried out with the original vectors, the cells transfected with this vector showed a peak containing EGFP positive cells in the flow histogram indicated with an arrow in Figure 3-19 and Figure

3-20. In the samples cotransfected with miRNA/siRNA the cell population defined as EGFP negative had grown remarkably and could conceivably contain cells with an EGFP expression down regulated to a level that would define them as EGFP negatives.

In summary, in this experiment there was a shift in the expression seen in the histograms; from cells transfected with the pEGFP-N3 vector, showing a peak of cells with high EGFP levels, to an increased number of EGFP negative cells when the cells were cotransfected with miRNA/siRNA. If the average intensity in the peak could be a measurement of EGFP expression of single cells in the experiment, and the shift could be a measure for miRNA/siRNA induced EGFP reduction, the precision of the flow cytometry experiments will be improved. Further experiments are required to assess if a change of promoter from PGK to CMV in the miRNA binding site reporter vectors will allow single cell measurement.



	-	siGFP-duplex
% EGFP-	29.0	54.3
% EGFP+	71.0	45.7

Figure 3-20: Representative histograms showed with reduced signal amplification, compared to the previous plots to look at the shape of the cell distributions. Theses plots can also be seen above in Figure 3-19. The fluorescence intensity (FITC-A) are plotted on a biexponential axis.

4 Discussion

MicroRNAs (miRNAs) are a relatively newly discovered class of small endogenous non coding RNA molecules with important roles in cell function and cancer development (Esquela-Kerscher and Slack, 2006; Mirnezami et al., 2009; Tong and Nemunaitis, 2008). miRNAs bind to complementary sequences of mRNA transcripts and cleave the mRNA transcripts or inhibit translation (Tong and Nemunaitis, 2008).

Grouping of tumor samples based on miRNA profiling has been shown to correlate with cell differentiation and development more accurately than mRNA expression profiling (Iorio et al., 2008), and analyses of miRNA expression have lead to suggestions for the use of miRNAs in cancer therapy (Esquela-Kerscher and Slack, 2006; Krutzfeldt et al., 2005; Meister et al., 2004; Meng et al., 2006; Slack and Weidhaas, 2006; Weiler et al., 2006).

The binding properties between the target mRNA sequence and the miRNA however, are complicated and diverse (Bartel, 2004, 2009; Brennecke et al., 2005; Grimson et al., 2007; Krek et al., 2005; Lewis et al., 2003; Sethupathy et al., 2006), and the levels of regulation range from weak to a full switch effect (Flynt and Lai, 2008). The problems regarding miRNA target site prediction emphasize the importance of the constructions of systems for functional verification of the miRNA target sequences.

In this thesis the aim was to construct a reporter vector system for cloning of 3'UTR sequences and with *EGFP* as the reporter gene. Flow cytometry was chosen as the method for analysing the EGFP expression of cells cotransfected with the constructed reporter vectors and predicted miRNAs (section 1.5).

4.1 The miRNA binding site reporter vector pmiRPG

The miRNA binding site reporter vector pmiRPG was constructed from pSiRPG, encoding the EGFP protein, through several cloning- and PCR steps and with two intermediate constructs, pRPG-1 and pRPG-2. The vector pmiRPG contains a functional multiple cloning

site in the 3'UTR sequence of the *EGFP* gene for cloning of 3'UTR sequences from selected genes.

The intermediates made in the vector construction process were verified with fragment analysis on agarose gels and with sequencing over the cloning junctions. The cloning junctions in the final construct pmiRPG were sequenced to verify the construct and to verify the orientation of the multiple cloning site (Figure 3-1 and Figure 3-3).

The three vector constructs pSiRPG, pmiRPG and pmiWWOX-3'U were analysed by flow cytometry after transient transfection into HEK293T cells. The results (Figure 3-9) showed that the three vectors displayed significant EGFP expression. This verified the reporter function of the pmiRPG constructs and indicated that cloning in the multiple cloning site of the pmiRPG vector did not disrupt the EGFP expression. The pmiRPG reporter vector constructed for expression studies, thereby seemed to be functional and suited for further experiments.

4.2 Verification of four 3'UTR constructs.

The three genes *WWOX, CXXC4 and WDR79,* were selected for the present study for their relations to cancer (section 1.4.2), and because they could potentially be candidates for miRNA regulation. *KRAS* was primarily selected as a positive control showing downregulation of the *KRAS* transcript by the let-7a miRNA (Johnson et al., 2005; Kanamoto et al., 2006). The 3'UTR sequences from *KRAS, WWOX and CXXC4* were amplified by PCR and cloned into the pmiRPG. The *WDR79* 3'UTR sequence was designed as an oligo nucleotide sequence and ligated directly into the vector. The vector constructs with 3'UTR sequences in pmiRPG as a fusion transcript with the *EGFP* gene, were sequenced over the cloning junctions to confirm the identity of each construct. The results from the sequencing analysis as shown in Figure 3-7 and in the Appendix, Figure 1-4, confirm that the four constructs contain the sequences from the transcripts *KRAS, WWOX, CXXC4* and *WDR79*. Two of these constructs, pmiKRAS-3'U and pmiWWOX-3'U were used in further experiments.

4.3 Use of flow cytometry to measure EGFP expression in transfected cells

Flow cytometry analysis was performed on HEK293T cells transiently transfected with the reporter constructs pSiRPG, pmiRPG, pmiKRAS-3'U, pmiWWOX-3'U and pEGFP-N3 encoding EGFP, in order to distinguish successfully transfected cells from non transfected cells. The purpose was to record fluorescence intensities from only the cells containing the vectors.

4.3.1 Flow cytometry detect decreased EGFP expression induced by siRNA

The down regulation of KRAS expression by the miRNA let-7a is well documented (Johnson et al., 2005; Kanamoto et al., 2006) and *KRAS* 3' UTR was included as a control in this thesis in order to verify the functionality of the vector system. To investigate whether the let-7a would give a detectable effect on the EGFP expression from the vector with a 3'UTR sequence from the *KRAS* gene, this construct was cotransfected with a synthetic let-7a oligo nucleotide into HEK293T cells.

In this experiment (Figure 3-10), the results showed that the EGFP expression from cells with pmiKRAS-3'U resulted in fluorescence intensities significantly above that of the control cells, not containing a vector. When the siGFP-duplex that targeted the *EGFP* transcript in the vector directly, was cotransfected in the control experiment, EGFP expression was reduced significantly. This experiment showed that flow cytometry was suited to detect the decreased EGFP expression induced by the siRNA.

When let-7a was cotransfected with the pmiKRAS-3'U in this experiment, a significant reduction in EGFP expression was observed (Figure 3-10). The *KRAS* transcript has previously been documented to be regulated by let-7 (Johnson et al., 2005; Kanamoto et al., 2006) and recent SNP (single nucleotide polymorphism) analysis of the let-7 target site in the *KRAS* 3'UTR have revealed an increase in lung cancer risk (Chin et al., 2008) and reduced survival in oral cancers (Christensen et al., 2009). These observations provide more evidences concerning the let-7 regulation of the KRAS expression. However, because

negative control siRNA or miRNA were not included in the experiment in the current thesis, the specificity of this regulation cannot be judged in this study.

4.3.2 All miRNAs tested result in decreased EGFP expression from pmiWWOX-3'U

An experiment was carried out to test the hypothesis that miRNAs predicted to bind to and regulate the *WWOX* gene, would give an effect on the EGFP expression from the reporter vector pmiWWOX-3'U. The four miRNAs let-7a, miR-516b*, miR-134 and miR-758 were cotransfected with pmiWWOX-3'U into HEK293T cells and analyzed with flow cytometry. The three latter miRNAs, were predicted by the miRNA target prediction programs to bind to the WWOX-3'UTR (section 3.3.3).

The results revealed down regulation of the EGFP expression from the pmiWWOX-3'UTR by all four miRNAs (Figure 3-11). The experiment may indicate that all four miRNAs regulate the *WWOX* gene, and there was no significant difference between the effect of the predicted and non predicted miRNA on the expression level. The results did however not rule out the possibility that the reduced level of expression recorded from the cotransfected cells could be caused by unspecific downregulation by the miRNAs.

A larger selection of miRNAs was tested in a new experiment, against cotransfecting with the same vector pmiWWOX-3'U. The miRNAs were added in two different concentrations to investigate if the wider selection of miRNAs would give the same downregulation as seen in the previous experiment, and to assay the effect of lowered miRNA concentration on the EGFP expression.

All the six miRNAs used in the experiment gave similar down regulation of the EGFP expression when 100pmol miRNA was used (Figure 3-12). The comparison between the two amounts of miRNAs tested revealed that one miRNA, miR-487a downregulated the EGFP expression similarly for both miRNA concentrations used. The remaining miRNAs showed less efficient downregulation when the miRNA concentration decreased. These results might indicate that the miRNA miR-487a specifically downregulated the transcript, while the other miRNAs tested gave unspecific down regulation of the expression with the highest

concentration, and lost much of the effect when the miRNA concentration was reduced. However the results showed that there were significant downregulations of the expression by all the six miRNAs added in both concentrations. The results when decreased miRNA concentrations were used, showing reduced repression effects might indicate that the decrease in the EGFP expression was caused by unspecific downregulation by the miRNAs.

The miR-487a gave interesting results concerning the regulation of pmiWWOX-3'U in this experiment, and more experiments needs to be done to establish this miRNA as a true regulator of WWOX expression.

4.3.3 Titration experiments indicate that improved signal recording methods are needed

It was hypothesized that titration experiments could reveal a more optimal concentration of miRNAs where the miRNAs specifically targeting a transcript would decrease the expression of the reporter vector and miRNAs that do not target the transcript would show limited effect on the expression. The pmiWWOX-3'U was chosen for the titration experiments, cotransfected with let-7a. This miRNA was not expected to target the *WWOX* 3'UTR since it had not been predicted by the five miRNA target prediction programs. The hypothesis was therefore that successful titrations with let-7a would lead to a concentration where it would fail to repress the EGFP expression from pmiWWOX-3'U.

Three separate titration experiments were completed with cotransfection of pmiWWOX-3'U and variable amounts of let-7a in attempts to reach the optimal amount of miRNAs, which would diminish unspecific regulation and show regulation of the EGFP expression from only the miRNAs that specifically target the 3'UTR.

The first experiment was carried out with four amounts of miRNAs in two fold dilutions from 100 – 12.5pmol and resulted in an increased EGFP expression at the lowest miRNA concentrations (Figure 3-13). Since all concentrations of miRNAs added in the experiment still repressed the EGFP expression, two new separate experiments were completed with four fold dilutions of miRNA down to 1.56pmol. The two experiments seen side by side in Figure 3-14 differed substantially and failed to reveal any tendency concerning the effect of the

titrated miRNA amounts even when the results were inspected separately. The results from the titration Experiment 1, in the figure, even suggested an up regulation of the signal from let-7a added in two amounts, 25 and 6.5pmol.

The results from these experiments could indicate that the method used to detect and present the expression levels in these experiments should be improved. The transfected cells did not show any distinct population (peak) apart from the untransfected cells in the flow experiments (representative histograms shown in Figure 3-16). Consequently the data used in this thesis are average intensity values from the cells showing fluorescence intensities that defines them as EGFP positive. This approach could allow influence from fluorescence background to affect the results. Transfected cells investigated under a microscope in UV-light (results not shown) indicated that the pmiRPG-based vectors transfected into HEK293T cells gave weak EGFP expression from the cells compared to other EGFP coding vectors such as pEGFP-N3. Most of the cells transfected with pmiRPG-based vectors and viewed under the microscope displaying weak EGFP expression, were defined as EGFP negatives in the flow cytometry experiments.

4.3.4 The CMV promoter effectively enhance the EGFP expression compared to a PGK promoter

The fluorescence intensity from the pmiRPG vectors with an *EGFP* gene expressed from a PGK promoter was compared against pEGFP-N3 with the *EGFP* gene expressed from the strong *CMV* promoter. The *EGFP* gene sequences in the two vector systems are identical.

The results showed an eight fold increase in the intensity from the pEGFP-N3 promoter compared to the three pmiRPG-based constructs (Figure 3-15). Comparisons of the flow histograms (representative histograms shown in Figure 3-16) revealed a difference in cell distributions in the sample with pEGFP-N3 transfected cells compared to cells transfected with pmiRPG-based constructs. The cells transfected with the PGK promoter based vector displayed low fluorescence intensities, and the dominant population of cells (70-90%) was found in the part of the histogram defined as EGFP negative. These cells displayed fluorescent intensities at the same level as untransfected cells. The cells transfected with the pEGFP-N3 vector in contrast, showed a population of over 80% of the cells in the area

defined as EGFP positive. A peak was also seen in the area with the highest value of fluorescence (above 10^5 FITC-A).

Considering that the four samples of cells were treated at the same time using the same protocol makes it tempting to assume that the transfection efficiency should be the same in the four cell samples. The reason why the distribution differs substantially between the two vector systems could therefore be that pSiRPG, pmiRPG and pmiWWOX-3'U must be transfected into each cell in higher quantities than pEGFP-N3 to make the cells display a fluorescence intensity above the normal auto fluorescence and to define the cells as EGFP positive. pEGFP-N3 might give the cells fluorescence properties even when only a few vector copies are present.

The observation from this experiment may suggest that the expression levels from the pmiRPG reporter vectors could be increased with an exchange of the promoter controlling the *EGFP* gene, to a stronger one, for example the CMV promoter explored in this experiment. A stronger expression could also reduce the influence from potential background noise and in that way improve the quality of the recordings.

4.3.5 Prolonged incubation after transfection leads to a marginal increase in EGFP expression

A hypothesis was made, that the recorded EGFP expression from the vectors could be increased if cells were analysed 48 hours after transfection. When the four vector constructs pSiRPG, pmiRPG, pmiWWOX-3'U and pEGFP-N3 were transfected into HEK293T (Figure 3-17), the EGFP levels were increased in the cells harvested after 48 hours. Even though the expression from the three PGK promoter based vectors showed increased expression after prolonged incubation, the expressions was still less than one third of the expression from pEGFP-N3 after 24 hours. In conclusion the increased time of incubation after transfection could not compensate for the less efficient EGFP expression from the PGK promoter. The CMV promoter seems to be more efficient and might be more suited for experiments where the fluorescence from EGFP is detected by flow cytometry. An exchange of the promoter in the pmiRPG constructs, could potentially give a better effect concerning the levels of the EGFP expression.

4.3.6 pEGFP-N3 transfected cells showed a “peak shift” after downregulation

A new experiment was designed to test the hypothesis that cells with stronger EGFP expression after transfection with the pEGFP-N3 vector would show a distinct peak in the flow histogram. Parallel experiments with pmiRPG, pmiWWOX-3’U and pEGFP-N3 cotransfection with miRNA/siRNA were performed to investigate whether expression from pEGFP-N3 would be downregulation in a similar way as seen with pmiRPG-based vectors in previous experiments (Figure 3-10, Figure 3-11).This experiment was carried out with transfection of the three vector constructs and cotransfection with miR-134. The siGFP-duplex used to downregulate the *EGFP* gene directly and the siGENOME Non-target siRNA#1 with no sequence similarity to the *EGFP* gene sequence, were used in parallel control experiments.

The results showed down regulation of the expression from all three vectors in all cotransfection experiments with miR-134, siGFP-duplex and siGENOME Non-Target siRNA#1(Figure 3-18). The CMV promoter containing vector, pEGFP-N3 seemed to be a subject for unspecific downregulation in the same extent as the other two vectors in the experiment, however one could speculate that new titration experiments with this vector would be more successful because of the stronger expression. The enhanced expression would have less influence from background fluorescence and might give more answers about the optimal miRNA concentration than obtained until now.

Histograms from the flow cytometry analysis of the pEGFP-N3 transfected cells (Figure 3-19) revealed a shift from high- to low fluorescence intensity in the flow cytometry histogram after downregulation. When two of the samples were recorded with lower signal amplification (Figure 3-20), a small peak with high fluorescence intensity was revealed in the histogram from cells transfected with pEGFP-N3, and the majority of cells appeared to shift towards low intensity after downregulation of the signal with the siGFP-duplex, seen as a taller peak containing EGFP negative cells.

The results obtained from this experiment suggested that stronger EGFP expression could make the analysis with flow cytometry easier to interpret. A distinct population of EGFP

positive cells shown by peaks in the histogram would make it easier to measure directly from this population and discard signals from other cells or background noise. If the peak shifts from high to low intensity in the plot after downregulation, this would be a direct visible effect in addition to the recorded values, and would facilitate using the average of the peak of EGFP positive cells directly as an estimate of EGFP expression in the transfected cells.

4.4 Conclusion

- ✓ A reporter construct pmiRPG was constructed, with *EGFP* as the reporter gene for incorporation of 3'UTR sequences as fusion transcript with the *EGFP* gene.
- ✓ 3'UTR sequences from the selected transcripts *KRAS, WWOX, WDR79 and CXXC4* were amplified or designed as oligos and cloned into the pmiRPG vector and verified with sequencing.
- ✓ The EGFP expression from the constructs was verified after cloning in the multiple cloning site of pmiRPG.
- ✓ All miRNAs/siRNAs, tested in this thesis, including a Non-Target siRNA, gave reduced EGFP expression. Further experiments need to be done to establish if this reduction results from unspecific regulation or inadequate recording methods.
- ✓ The hsa-miR-487 gave reduced EGFP recordings with reduced miRNA concentrations in an experiment where hsa-miR-487 and reporter vector pmiWWOX-3'U were cotransfected into HEK293T cells. Additional experiments are required to establish this miRNA as a true regulator of WWOX expression.
- ✓ The CMV promoter was shown to give increased EGFP expression compared to the vectors with PGK promoter. The CMV promoter could be inserted in the pmiRPG vector to enhance EGFP expression.

4.4.1 Suggestions for further work

The results presented in this study indicate that the flow cytometry recordings after an exchange of promoter from a PGK- to a CMV-promoter in the pmiRPG might facilitate using the "peak shift" as seen in Figure 3-20 as an estimate of the miRNA knock down. This could be a better measure of knock down than the average values of intensities used in the experiments in this thesis. If a promoter exchange and new titration experiments gave satisfying results, the miRNA target prediction verification system could prove its value. Some important points could be followed in the extension of this thesis.

- Exchange of the promoter driving the expression of the reporter gene *EGFP* in the pmiRPG vectors from PGK to CMV to achieve higher fluorescence intensities in flow cytometry experiments. This might result in distinct peaks of the EGFP positive cells and result in peak shifts during a downregulation of the signal.

- Make stable transfected cell lines with pmiKRAS-3'U and pmiWWOX-3'U for titration experiments with let-7a in an attempt to reach an optimal miRNA concentration where only specific regulation influence the EGFP expression.

- Use stably transfected cell lines for further experiments analyzing the effect of hsa-miR-487 on the expression levels of pmiWWOX-3'U to test whether this miRNA is a true regulator of the WWOX transcript.

References

Adjei, A.A. (2001). Blocking Oncogenic Ras Signaling for Cancer Therapy. J Natl Cancer Inst *93*, 1062-1074.

Aqeilan, R.I., and Croce, C.M. (2007). WWOX in biological control and tumorigenesis. J Cell Physiol *212*, 307-310.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Bednarek, A.K., Laflin, K.J., Daniel, R.L., Liao, Q., Hawkins, K.A., and Aldaz, C.M. (2000). WWOX, a novel WW domain-containing protein mapping to human chromosome 16q23.3-24.1, a region frequently affected in breast cancer. Cancer Res *60*, 2140-2145.

Bottoni, A., Zatelli, M.C., Ferracin, M., Tagliati, F., Piccin, D., Vignali, C., Calin, G.A., Negrini, M., Croce, C.M., and Degli Uberti, E.C. (2007). Identification of differentially expressed microRNAs by microarray: a possible role for microRNA genes in pituitary adenomas. J Cell Physiol *210*, 370-377.

Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. (2005). Principles of microRNA-target recognition. PLoS Biol *3*, e85.

Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., *et al.* (2002). Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A *99*, 15524-15529.

Calin, G.A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S.E., Iorio, M.V., Visone, R., Sever, N.I., Fabbri, M., *et al.* (2005). A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med *353*, 1793-1801.

Calin, G.A., Sevignani, C., Dumitru, C.D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., *et al.* (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A *101*, 2999-3004.

Chin, L.J., Ratner, E., Leng, S., Zhai, R., Nallur, S., Babar, I., Muller, R.-U., Straka, E., Su, L., Burki, E.A., *et al.* (2008). A SNP in a let-7 microRNA Complementary Site in the KRAS

3' Untranslated Region Increases Non-Small Cell Lung Cancer Risk. Cancer Res *68*, 8535-8540.

Christensen, B.C., Moyer, B.J., Avissar, M., Ouellet, L.G., Plaza, S., McClean, M.D., Marsit, C.J., and Kelsey, K.T. (2009). A let-7 microRNA binding site polymorphism in the KRAS 3' UTR is associated with reduced survival in oral cancers. Carcinogenesis, bgp099.

Cohen, S.N., Chang, A.C., and Hsu, L. (1972). Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. Proc Natl Acad Sci U S A *69*, 2110-2114.

Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer *3*, 11-22.

Driouch, K., Prydz, H., Monese, R., Johansen, H., Lidereau, R., and Frengen, E. (2002). Alternative transcripts of the candidate tumor suppressor gene, WWOX, are expressed at high levels in human breast tumors. Oncogene *21*, 1832-1840.

Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D.S. (2003). MicroRNA targets in Drosophila. Genome Biol *5*, R1.

Esquela-Kerscher, A., and Slack, F.J. (2006). Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer *6*, 259-269.

Flynt, A.S., and Lai, E.C. (2008). Biological principles of microRNA-mediated regulation: shared themes amid diversity. Nat Rev Genet *9*, 831-842.

Garcia-Closas, M., Kristensen, V., Langerod, A., Qi, Y., Yeager, M., Burdett, L., Welch, R., Lissowska, J., Peplonska, B., Brinton, L., *et al.* (2007). Common genetic variation in TP53 and its flanking genes, WDR79 and ATP1B2, and susceptibility to breast cancer. Int J Cancer *121*, 2532-2538.

Giehl, K. (2005). Oncogenic Ras in tumour progression and metastasis. Biological Chemistry *386*, 193-205.

Griffiths-Jones, S. (2004). The microRNA Registry. Nucl Acids Res *32*, D109-111.

Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., and Enright, A.J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. Nucl Acids Res *34*, D140-144.

Griffiths-Jones, S., Saini, H.K., van Dongen, S., and Enright, A.J. (2008). miRBase: tools for microRNA genomics. Nucleic Acids Res *36*, D154-158.

Grimson, A., Farh, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell *27*, 91-105.

Hanahan, D., and Weinberg, R.A. (2000). The Hallmarks of Cancer. Cell *100*, 57-70.

Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y., and Takahashi, T. (2005). A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res *65*, 9628-9632.

Heim, R., Cubitt, A.B., and Tsien, R.Y. (1995). Improved green fluorescence. Nature *373*, 663-664.

Herzenberg, L.A., Tung, J., Moore, W.A., and Parks, D.R. (2006). Interpreting flow cytometry data: a guide for the perplexed. Nat Immunol *7*, 681-685.

Hezova, R., Ehrmann, J., and Kolar, Z. (2007). WWOX, a new potential tumor suppressor gene. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub *151*, 11-15.

Hino, S., Kishida, S., Michiue, T., Fukui, A., Sakamoto, I., Takada, S., Asashima, M., and Kikuchi, A. (2001). Inhibition of the Wnt signaling pathway by Idax, a novel Dvl-binding protein. Mol Cell Biol *21*, 330-342.

Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. Gene *96*, 23-28.

Iorio, M.V., Casalini, P., Tagliabue, E., Menard, S., and Croce, C.M. (2008). MicroRNA profiling as a tool to understand prognosis, therapy response and resistance in breast cancer. Eur J Cancer *44*, 2753-2759.

Iorio, M.V., Ferracin, M., Liu, C.G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., *et al.* (2005). MicroRNA gene expression deregulation in human breast cancer. Cancer Res *65*, 7065-7070.

John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human MicroRNA targets. PLoS Biol *2*, e363.

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. (2005). RAS is regulated by the let-7 microRNA family. Cell *120*, 635-647.

Kanamoto, T., Terada, K., Yoshikawa, H., and Furukawa, T. (2006). Cloning and regulation of the vertebrate homologue of lin-41 that functions as a heterochronic gene in Caenorhabditis elegans. Dev Dyn *235*, 1142-1149.

Kim, V.N., and Nam, J.W. (2006). Genomics of microRNA. Trends Genet *22*, 165-173.

Kiriakidou, M., Nelson, P.T., Kouranov, A., Fitziev, P., Bouyioukos, C., Mourelatos, Z., and Hatzigeorgiou, A. (2004). A combined computational-experimental approach predicts human microRNA targets. Genes Dev *18*, 1165-1178.

Kojima, T., Shimazui, T., Hinotsu, S., Joraku, A., Oikawa, T., Kawai, K., Horie, R., Suzuki, H., Nagashima, R., Yoshikawa, K., *et al.* (2009). Decreased expression of CXXC4 promotes a malignant phenotype in renal cell carcinoma by activating Wnt signaling. Oncogene *28*, 297-305.

Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., *et al.* (2005). Combinatorial microRNA target predictions. Nat Genet *37*, 495-500.

Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K.G., Tuschl, T., Manoharan, M., and Stoffel, M. (2005). Silencing of microRNAs in vivo with 'antagomirs'. Nature *438*, 685-689.

Laitinen, J., Rakkolainen, T., and Holtta, E. (2004). Vectorettes for long chromosome walking in genomic DNA of the human p53 gene. Biotechniques *37*, 674-676, 678.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. Cell *120*, 15-20.

Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798.

Liu, C.G., Calin, G.A., Meloon, B., Gamliel, N., Sevignani, C., Ferracin, M., Dumitru, C.D., Shimizu, M., Zupo, S., Dono, M., *et al.* (2004). An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc Natl Acad Sci U S A *101*, 9740-9744.

Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., *et al.* (2005). MicroRNA expression profiles classify human cancers. Nature *435*, 834-838.

Ludes-Meyers, J.H., Bednarek, A.K., Popescu, N.C., Bedford, M., and Aldaz, C.M. (2003). WWOX, the common chromosomal fragile site, FRA16D, cancer gene. Cytogenet Genome Res *100*, 101-110.

Malumbres, M., and Barbacid, M. (2003). RAS oncogenes: the first 30 years. Nat Rev Cancer *3*, 459-465.

Marques, S.M., and Esteves da Silva, J.C. (2009). Firefly bioluminescence: a mechanistic approach of luciferase catalyzed reactions. IUBMB Life *61*, 6-17.

Mattila, P., Korpela, J., Tenkanen, T., and Pitkanen, K. (1991). Fidelity of DNA synthesis by the Thermococcus litoralis DNA polymerase--an extremely heat stable enzyme with proofreading activity. Nucleic Acids Res *19*, 4967-4973.

Meister, G., Landthaler, M., Dorsett, Y., and Tuschl, T. (2004). Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. RNA *10*, 544-550.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature *431*, 343-349.

Meng, F., Henson, R., Lang, M., Wehbe, H., Maheshwari, S., Mendell, J.T., Jiang, J., Schmittgen, T.D., and Patel, T. (2006). Involvement of human micro-RNA in growth and response to chemotherapy in human cholangiocarcinoma cell lines. Gastroenterology *130*, 2113-2129.

Mirnezami, A.H., Pickard, K., Zhang, L., Primrose, J.N., and Packham, G. (2009). MicroRNAs: key players in carcinogenesis and novel therapeutic targets. Eur J Surg Oncol *35*, 339-347.

Muller, S.R., Sullivan, P.D., Clegg, D.O., and Feinstein, S.C. (1990). Efficient Transfection and Expression of Heterologous Genes in PC12 Cells. DNA and Cell Biology *9*, 221-229.

Negrini, M., Ferracin, M., Sabbioni, S., and Croce, C.M. (2007). MicroRNAs in human cancer: from research to therapy. J Cell Sci *120*, 1833-1840.

Nolan, J.P., and Yang, L. (2007). The flow of cytometry into systems biology. Brief Funct Genomic Proteomic *6*, 81-90.

O'Donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V., and Mendell, J.T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. Nature *435*, 839-843.

Ota, A., Tagawa, H., Karnan, S., Tsuzuki, S., Karpas, A., Kira, S., Yoshida, Y., and Seto, M. (2004). Identification and Characterization of a Novel Gene, C13orf25, as a Target for 13q31-q32 Amplification in Malignant Lymphoma. Cancer Res *64*, 3087-3095.

Patterson, G.H., Knobel, S.M., Sharif, W.D., Kain, S.R., and Piston, D.W. (1997). Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. Biophysical Journal *73*, 2782-2790.

Pimiento, J.M., Teso, D., Malkan, A., Dudrick, S.J., and Palesty, J.A. (2007). Cancer of unknown primary origin: a decade of experience in a community-based hospital. Am J Surg *194*, 833-837; discussion 837-838.

Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., and Cormier, M.J. (1992). Primary structure of the Aequorea victoria green-fluorescent protein. Gene *111*, 229-233.

Quinlan, M.P., and Settleman, J. (2009). Isoform-specific ras functions in development and cancer. Future Oncology *5*, 105-116.

Rajalingam, K., Schreck, R., Rapp, U.R., and Albert, S. (2007). Ras oncogenes and their downstream targets. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1773*, 1177-1195.

Rajewsky, N. (2006). microRNA target predictions in animals. Nat Genet *38 Suppl*, S8-13.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-906.

Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. Genome Res *14*, 1902-1910.

Rosenfeld, N., Aharonov, R., Meiri, E., Rosenwald, S., Spector, Y., Zepeniuk, M., Benjamin, H., Shabes, N., Tabak, S., Levy, A., *et al.* (2008). MicroRNAs accurately identify cancer tissue origin. Nat Biotechnol *26*, 462-469.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988). Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase. Science *239*, 487-491.

Salzman, D.W., Shubert-Coleman, J., and Furneaux, H. (2007). P68 RNA Helicase Unwinds the Human let-7 MicroRNA Precursor Duplex and Is Required for let-7-directed Silencing of Gene Expression. J Biol Chem *282*, 32773-32779.

Sambrook, J., and Russel, D., W., eds. (2001). Molecular Cloning, 3 edn (New York, Cold Spring Harbour Laboratory Press).

Sethupathy, P., Megraw, M., and Hatzigeorgiou, A.G. (2006). A guide through present computational approaches for the identification of mammalian microRNA targets. Nat Methods *3*, 881-886.

Slack, F.J., and Weidhaas, J.B. (2006). MicroRNAs as a potential magic bullet in cancer. Future Oncology *2*, 73-82.

Smith, T.F., Gaitatzes, C., Saxena, K., and Neer, E.J. (1999). The WD repeat: a common architecture for diverse functions. Trends in Biochemical Sciences *24*, 181-185.

Størvold, G.L., Gjernes, E., Askautrud, H.A., Borresen-Dale, A.L., Perou, C.M., and Frengen, E. (2007). A retroviral vector for siRNA expression in mammalian cells. Mol Biotechnol *35*, 275-282.

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., *et al.* (2004). Reduced Expression of the let-7 MicroRNAs in Human Lung Cancers in Association with Shortened Postoperative Survival. Cancer Res *64*, 3753-3756.

Tang, G. (2005). siRNA and miRNA: an insight into RISCs. Trends in Biochemical Sciences *30*, 106-114.

Tong, A.W., and Nemunaitis, J. (2008). Modulation of miRNA activity in human cancer: a new paradigm for cancer gene therapy? Cancer Gene Ther *15*, 341-355.

Tycowski, K.T., Shu, M.-D., Kukoyi, A., and Steitz, J.A. (2009). A Conserved WD40 Protein Binds the Cajal Body Localization Signal of scaRNP Particles. Molecular Cell *34*, 47-57.

Volinia, S., Calin, G.A., Liu, C.G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., *et al.* (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A *103*, 2257-2261.

Watanabe, Y., Tomita, M., and Kanai, A. (2007). Computational methods for microRNA target prediction. Methods Enzymol *427*, 65-86.

Weidhaas, J.B., Babar, I., Nallur, S.M., Trang, P., Roush, S., Boehm, M., Gillespie, E., and Slack, F.J. (2007). MicroRNAs as potential agents to alter resistance to cytotoxic anticancer therapy. Cancer Res *67*, 11111-11116.

Weiler, J., Hunziker, J., and Hall, J. (2006). Anti-miRNA oligonucleotides (AMOs): ammunition to target miRNAs implicated in human disease? Gene Ther *13*, 496-502.

Yanaihara, N., Caplen, N., Bowman, E., Seike, M., Kumamoto, K., Yi, M., Stephens, R.M., Okamoto, A., Yokota, J., Tanaka, T., *et al.* (2006). Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell *9*, 189-198.

Zhang, B., Pan, X., Cobb, G.P., and Anderson, T.A. (2007). microRNAs as oncogenes and tumor suppressors. Dev Biol *302*, 1-12.

Appendix

Table 1: Primer sequences

Primers	Sequence 5'-3'
pBABEp4-Fwr	agt tct tgc agc tcg gtg ac
pSiRPG-7166-Rev	tga aag ctg cgc act agt gag gga gct cat agg ccg gca tag acg cgt ttc ttc tga cac aac agt ctc
pSiRPG-7167-Fwr	acg cgt cta tgc cgg cct atg agc tcc ctc act agt gcg cag ctt tca tca agc tga tct gag tcc gg
pSRPG-Rev	gtc tct ccc cct tga acc tc
pmiRPG-5850-Fwr	gac ttt cca cac ctg gtt gc
pmiRPG-6136-Rev	aca tgg tcc tgc tgg agt tc
WWOX-3'UTR-Fwr	aatt acg cgt gtg gag ctc aga gcg gat g
WWOX-3'UTR-Rev	aatt gcg gcc gcg ata ttc aag cat gtt ctt tta tta ag
KRAS-3'UTR-Fwr	aatt acg cgt gtt aca cca tct tca gtg cca g
KRAS-3'UTR-Rev	aatt gcg gcc gcc aca taa agg taa tta acc act ac
CXXC4-3'UTR-Fwr	aatt acg cgt gaa gca ttc cga tgg ttc tt
CXXC4-3'UTR-Rev	aatt gcg gcc gct tga aga aat ccc tat aca taa aat ga

Table 2: Oligo nucleotide linker sequences

Oligo sequences	Sequence 5'-3'
MABNB-1	cgc gtg ggc cca gat ctg cgg ccg cg
BNBAM-2	cgc gcg cgg ccg cag atc tgg gcc ca
WDR79-3'UTR-1	cgc gtt ata aaa agg ttt tta tga tac tag agc
WDR79-3'UTR-2	ggc cgc tct agt atc ata aaa acc ttt tta taa

Table 3: siRNA sequence/target sequence

siRNA	Sequence 5'-3'
siGFP-duplex	ggc tac gtc cag gag cgc acc (Target seq.)
Non-Target siRNA#1	uag cga cua aac aca uca auu

Table 4: Mature miRNA sequences

miRNA Stem-loop	miRNA Mature	miRNA-Mature Sequence 5'-3'
hsa-let-7a-1	hsa-let-7a	uga ggu agu agg uug uau agu u
hsa-mir-758	hsa-miR-758	uuu gug acc ugg ucc acu aac c
hsa-mir-516b-1	hsa-miR-516b*	ugc uuc cuu uca gag ggu
hsa-mir-134	hsa-miR-134	ugu gac ugg uug acc aga ggg g
hsa-mir-205	hsa-miR-205	ucc uuc auu cca ccg gag ucu g
hsa-mir-570	hsa-miR-570	cga aaa cag caa uua ccu uug c
hsa-miR-105-2	hsa-miR-105	uca aau gcu cag acu ccu gug gu
hsa-miR-532	hsa-miR-532-5p	cau gcc uug agu gua gga ccg u
hsa-mir-487a	hsa-miR-487a	aauc aua cag gga cau cca guu
hsa-miR-548d-2	hsa-miR-548d-3p	caa aaa cca cag uuu cuu uug c

Figure 1: pmiKRAS-3'UTR aligned against KRAS

The aligned sequences correspond to the sequenced pmiKRAS-3'U and the *KRAS* transcript (GenBank: M54968). The sequencing result is presented in bold, and the restriction sites are underlined.

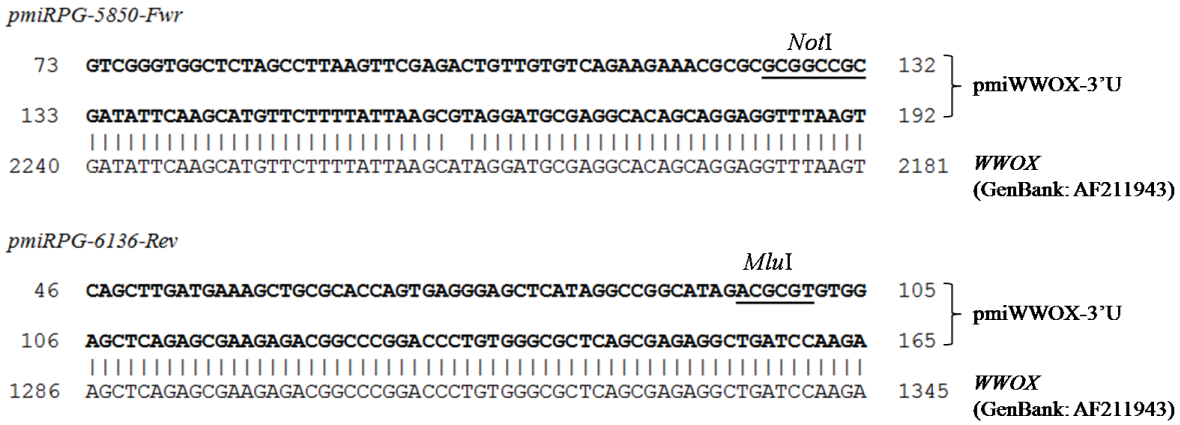
Figure 2: pmiCXXC4-3'UTR aligned against CXXC4

The aligned sequences correspond to the sequenced pmiCXXC4-3'U and the *CXXC4* transcript (GenBank: NM_025212). The sequencing result is presented in bold, and the restriction sites are underlined.

Figure 3: pmiWDR79-3'UTR aligned against WDR79

The aligned sequences correspond to the sequenced pmiWDR79-3'U and the *WDR79* transcript (GenBank: NM_018081). The sequencing result is presented in bold, and the restriction sites are underlined.

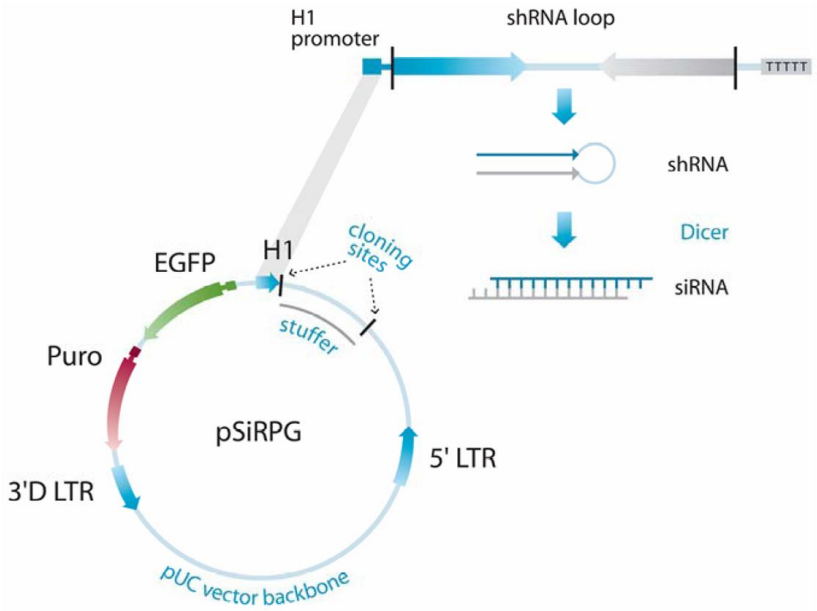
Figure 4: pmiWWOX-3'UTR aligned against WWOX



The aligned sequences correspond to the sequenced pmiWWOX-3'U and the *WWOX* transcript (GenBank: AF211943). The sequencing result is presented in bold, and the restriction sites are underlined.

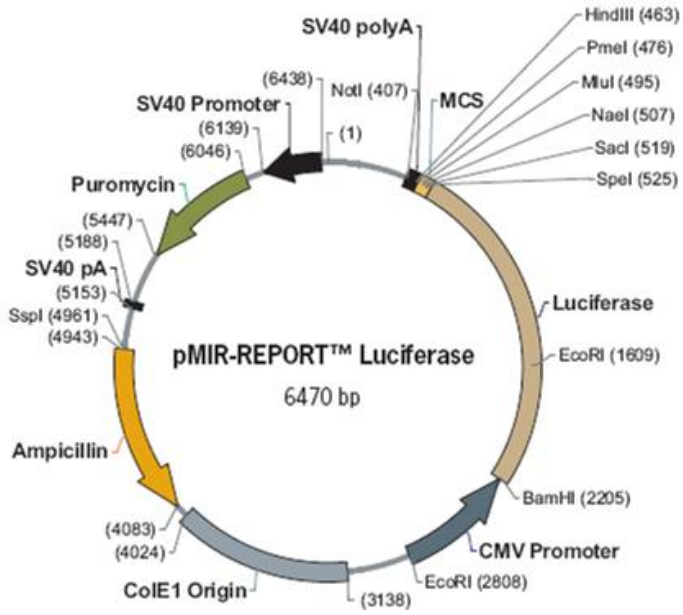
Vector constructs used directly or indirectly in the thesis

Figure 5: pSiRPG



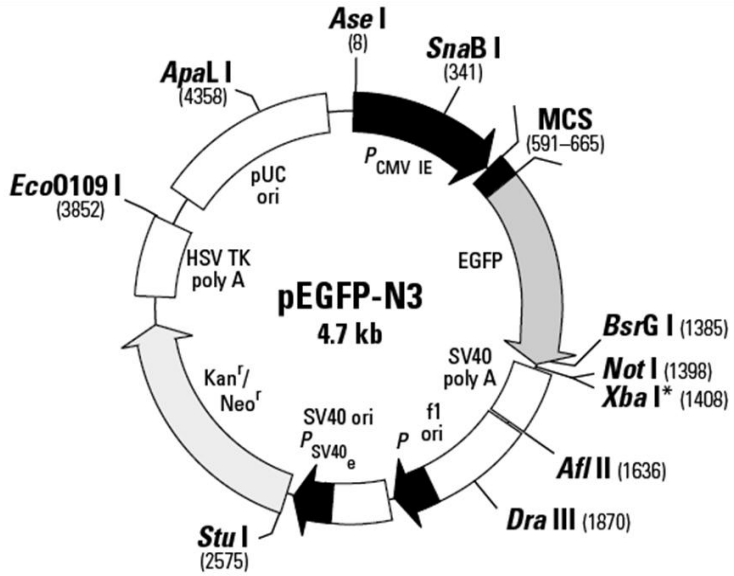
The expression vector pSiRPG (Størvold et al., 2007) was used as the vector backbone in the construction of pmiRPG.

Figure 6: pMIR-REPORT Luciferase



A sequence from the luciferase based reporter vector pMIR-REPORT (Applied Biosystems Inc, Foster City, CA, USA) was cloned into pmiRPG during one of the cloning steps.

Figure 7: pEGFP-N3 (GenBank U57609)



The construct pEGFP-N3 (Becton Dickinson, Rutherford, NJ,USA) has an *EGFP* gene behind the immediate early promoter CMV, and was used in experiments to study the fluorescence intensity from expressed EGFP in transfected cells.